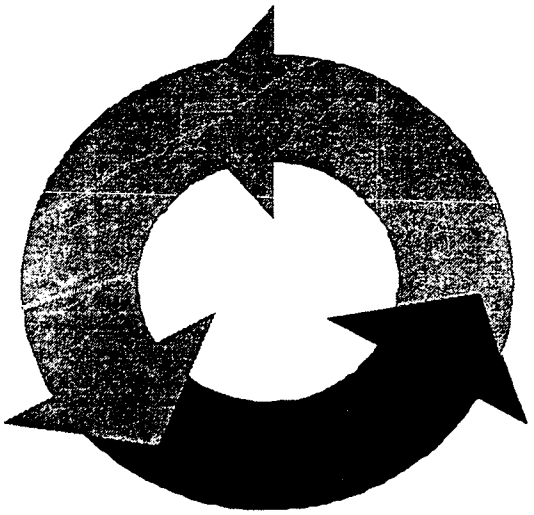
DOCUMENT RESUME

ED 429 112 TM 029 654

TITLE	Classroom Assessment: Linking Instruction and Assessment [and] Study Guide.
INSTITUTION	North Carolina State Dept. of Public Instruction, Raleigh.
PUB DATE	1999-01-00
NOTE	194p.
PUB TYPE	Guides - Non-Classroom (055)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	*Educational Assessment; Elementary Education; *Instructional Effectiveness; *State Programs; Teaching Methods; Test Construction; Test Use; *Testing Programs
IDENTIFIERS	*North Carolina

ABSTRACT

The bond that links quality assessment and effective teaching is clarified for North Carolina teachers in the context of the state's continued emphasis on effective schooling. This manual begins a statewide effort to promote assessment literacy among educators. It is emphasized that no single type of assessment can meet all of the purposes of assessment or information needs of the various educational decision makers. The assessment cycle model of assessment on which the discussion is based highlights the interplay among learning targets as defined in the North Carolina "Standard Course of Study," assessment methods, and decisions and actions. The following chapters contain specific information about: (1) "Classroom Assessment and Instruction"; (2) "Clarifying Learning Targets"; (3) "Using Multiple Assessment Strategies"; (4) "Making Decisions and Taking Action"; and (5) "Documenting and Communicating." Contains glossary. The study guide is included. (Contains one table and six references.) (SLD)



Classroom Assessment:

Linking Instruction and Assessment

Public Schools of North Carolina
State Board of Education
North Carolina Department of Public Instruction
Raleigh, North Carolina

January 1999

Thanks to the following Classroom Assessment Committee members who were instrumental in producing this document:

Kymm Ballard	Instructional Services
Mildred Bazemore	Accountability Services
Belinda Black	Accountability Services
Dee Brewer	Accountability Services
Carolyn Cobb	Accountability Services
Bryar Cougle	Instructional Services
Jane Cowan	Instructional Services
Jennie Franklin	School Improvement
Charlotte Hughes	School Improvement
Jeane Joyner	Accountability Services
Barbara Kuligowski	Instructional Services
Frances Melott	School Improvement
Daisy Vickers	Accountability Services
Rhonda Welfare	Instructional Services
Patricia Wilkins	School Improvement
Jan Williamson	Instructional Services
Gary Williamson	Accountability Services

If you have questions or comments on these materials, please contact Jeane Joyner (919-715-1864) or Jan Williamson (919-715-1875).

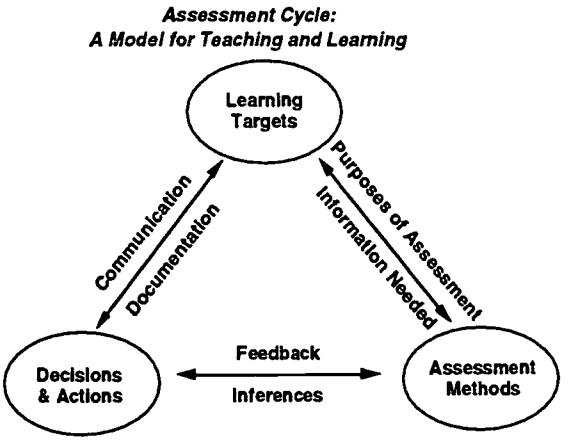


Table of Contents

Overview

Understanding Assessment in North Carolina: The Role of Assessment in Teaching and Learning

The Current Assessment Climate

Classroom assessment supports testing for accountability.

The current climate of high-stakes accountability in North Carolina and throughout the nation naturally focuses school administrators and teachers on assessments that measure student achievement for accountability purposes. While classroom assessment has always been an important element of instruction, it takes on a special urgency in today's high stakes environment. In this environment, many teachers are tempted to de-emphasize day-to-day assessment to focus on assessment tools more obviously linked to accountability measures - in this case North Carolina's End-of-Grade and End-of-Course Tests. When this shift occurs, classroom assessment becomes narrowed and looks more like multiple choice (and other selected choice) tests used at the state level. Instruction and classroom activities also become limited to only those parts of the curriculum that are tested. Writing instruction, for example, becomes narrowed in an attempt to mirror the writing prompts used on writing tests at grades 4, 7, and 10.

It is easy to understand why the high-stakes consequences of our ABCs Accountability Model lead many school staff to believe that it is essential to focus on assessments that resemble state assessments. The thinking goes, "If we do not teach and test in the classroom the way the state tests, we will not make expected growth on the ABCs." However, this thinking limits the learning opportunities for students, the likelihood of mastering the North Carolina *Standard Course of Study* goals and objectives, and the rich possibilities for using assessment as a teaching tool. While schools may well achieve expected or exemplary growth by narrowing the teaching and testing focus to resemble state assessments, this approach is short-sighted. In addition, there are many schools making exemplary growth that are *not* limiting teaching and learning in this way.

The North Carolina Department of Public Instruction supports the continued development of quality teaching and classroom assessment as vehicles to prepare students to master rigorous content and performance standards as well as to do well on accountability measures. We believe that the strategies most likely to result in long-term growth and learning of high quality will result from effective use of quality classroom assessments as an integral part of instruction. Additionally, strong classroom assessment engages students in self-assessment and greater ownership of their own learning.

"Evaluation" comes from the Latin meaning "to value." "Assessment" comes from the Latin meaning "to sit beside."

Thus, this document is designed to clarify further the bond that links quality assessment and effective teaching - and subsequently effective schools. Quality classroom assessment is essential to the goals of high student achievement and continuous improvement of schools. Learning takes place one student at a time, and quality teaching-which includes quality assessment-is essential to reaching the high stakes accountability goals.

Assessment Literacy

At one time "being literate" meant being able to sign one's name. It gradually grew to mean being able to read and write. As the skills necessary to be successful in an increasingly complex society grew, the meaning of literacy has grown as well. The "new literacy" goes beyond the basics to include *capacities once demanded of a privileged...elite: to think critically and creatively, solve problems, exercise judgment, and learn new skills and knowledge throughout a lifetime* (Brown, 1991). How does this relate to assessment literacy?

Quality teaching – which includes quality classroom assessment – is essential to reaching the high stakes accountability goals.

Using assessment effectively begins with **assessment literacy**: understanding sound principles of assessment, understanding that there are different types of assessments for different purposes and levels of education, being able to choose appropriate assessment methods for designated learning goals and targets, and using assessment to motivate and teach students. Sound principles of assessment apply to classroom assessment as well as high-stakes accountability assessment. However, they play out differently. Most educators do not get this kind of "assessment education" in college; it is usually left out or left for on-the-job-training. When they are not assessment-literate, educators may choose assessment methods

inappropriate for their purposes. There are essential assessment principles and concepts that can help teachers to choose effective methods of assessing learning and even to use those methods as instructional tools.

Why should we care? Because assessment is an integral part of instruction. Good teaching involves good assessing. This manual begins a statewide initiative to promote assessment literacy among educators in North Carolina. We hope it will provide you with the essentials of assessment literacy and stimulate you to think further about how to use assessment creatively and effectively to promote high levels of learning in your classroom - and, by the way - to score well on state accountability assessments.

Good teaching involves good assessment.

What is Assessment?

Assessment often means different things to different people. When groups of educators were asked what they think of when they hear the word "assessment," responses ranged widely. Responses included multiple choice tests, essays, making judgments, End-of-Grade/Course Tests, paper-and-pencil, "stress," authentic assessment, observations, portfolios, journals. Assessment can be, at times, any of these things.

However, it is important to look at a common definition of assessment to guide assessment selection and design. In one sense, assessment is simply "gathering information to make decisions." But the type of decision to be made (i.e., the purpose of the assessment) guides the types of information that needs to be gathered. More formally, assessment means:

> a sample of student performance or behavior used to obtain information or provide feedback about learning targets in order to make decisions and take action about individual students, groups of students, instruction, programs or schools.

The learning targets for students in North Carolina are identified in the *Standard Course of Study*.

Let's examine the meaning of "sample of performance."

Sample of Student Performance: Whether teachers are carrying out classroom assessments specific to their learning targets or the state is conducting the End-of-Grade Test, any

assessment method collects a *sample* of student performance. Rarely can we assess the actual end performance that we are working toward for application in the everyday world. Rather, we sample learning in various ways in order to make inferences about how well students are mastering and are able to use knowledge and skills. All assessments rely on a relatively small number of exercises to permit the user to draw inferences about a student's mastery of larger domains of achievement. In the classroom, teachers use a variety of ways to collect performance samples: e.g., multiple choice tests for content knowledge, open-ended items for application and reasoning, real-life extended problems for a closer assessment of students' ability to apply these skills outside the classroom. In fact, when assessments approximate real-life applications, they may be called "authentic" or direct assessments. These performance samples are much broader and involve more time. The closer we get to the classroom, the more samples of behavior or performances we have in order to make accurate judgments about a student's overall learning in a given domain.

The number and type of samples of performance required for accurate assessment also relate to the **purposes** of assessment; that is, the types of decisions that are to be made as a result of this assessment. It is also a balance between having an adequate number of samples to give a complete picture of learning and efficiency of time and effort. For example, state assessments are intended for accountability purposes, meaning decisions about how well students in schools are learning broadly defined objectives. For this purpose, state and local policy makers typically require only a "snapshot" of student achievement in a content domain - a small sample of learning. Thus, a multiple-choice test is an efficient way to get a sample of each student's performance. This sample, however, is inadequate for a teacher to judge true student mastery of a learning target or content domain; it is too limited a sample of student performance.

Teachers need a lot of information about each student.

Policy makers need samples of information about a lot of students.

In short, teachers typically need several samples of performance on *each student* **for classroom instructional decisions.** Accountability assessments need a limited sample of performance on *many students*.

Purposes of Assessment

Another key concept in assessment literacy is that different

purposes of assessment call for different types of assessments. Several purposes of assessment can be identified and are linked to who is asking the questions about learning and who will use the information. Purpose also often relates to the level of education (e.g., state, district/central office, school, classroom). Selected purposes are discussed here. Table 1 on the next page illustrates these purposes, the primary users or persons asking the questions, and typical questions they ask. The types of assessments may or may not always be different across different purposes; however, they likely will be used differently based on the purpose of the assessment.

• Accountability. Policy makers at the state and local level usually ask questions about learning for accountability purposes. Their questions usually are targeted at schools or districts, not at individual students. They typically are concerned about overall achievement and learning rather than specific learning targets. They are accountable to the public and are asking if the public is getting a return on its investment and if students are meeting specified standards. Their learning questions are broad (e.g., Are students learning the "basics?") and, in traditional accountability systems, require only a small sample of student performance to answer those questions. More recent proposals for student accountability (e.g., high school exit criteria or promotion standards) may require more frequent and diverse samples of performance or more extended kinds of performances (e.g., senior exit project).

Assessment for accountability gives a picture of overall achievement and learning.

• Program Evaluation. Program evaluation can occur at the state, district, or school level. Users typically are educators who must make decisions about how to improve programs, where program weaknesses and strengths occur, and what programs should be expanded or discontinued. Depending on the types of questions, different assessment **methods** could be used. However, principals might use state EOG/EOC Tests to determine if students are learning more in mathematics than reading, or if certain grade levels consistently do better than others. Or analysis of specific types of student work across many students might be carried out for greater precision in assessment of specific learning targets. In these instances, program evaluation merges with instructional leadership (see Table 1), because the principal and school staff may use these data to determine needed professional development or changes in school curriculum.

Table 1 Examples of Links Between Purposes and Methods of Assessment *

Purpose	Primary Users	Typical Questions	Type of Information Needed	Possible Assessment Method(s)
Accountability	• State & /Local Policy Makers • Public/Community	1. Are schools producing results in the basics? 2. Are students learning the *Standard Course of Study?*	Periodic Assessment of Group Achievement (typically once/year)	Multiple choice tests Possibly some performance tasks (Writing)
Program Evaluation	• Superintendents/Local Boards • Principals • State Policy Makers	1. Are our programs producing student learning? 2. Which schools need more assistance?	Periodic Assessment of Group Achievement	Multiple choice tests Performance tests/tasks
Instructional Leadership	• School Administrators	1. Are teachers and instructional strategies in given areas producing results? 2. What kinds of professional development would help? 3. How shall we spend building resources in support of instruction? 4. What does this teacher need to assure student competence?	Periodic Assessment of Group Achievement Examination of Student Work (synthesis of group results) Continuous Assessment of Group Achievement	Multiple choice Possible (more difficult to summarize): Open-ended tests, performance tasks, ongoing student work (portfolios)
Instruction: Classroom	• Teachers	1. Are my teaching strategies working? 2. What do *these students* need help with? 3. What do students understand; what can they apply?	Continuous Assessment of Group Performance/Achievement Continuous Assessment of Individual Performance Summarized Over Group	Multiple means to include multiple choice, open-ended, performance Multiple means as above, plus observation and class discussion
Instruction & Diagnosis: Individual	• Teachers	1. What does *this student* need help with? 2. What misconceptions/strengths does this student have?	Continuous Assessment of Individual Mastery/Performance	Multiple means to include observation, conversations, analysis of student work

* There are other purposes and users of assessment information. These are samples only.

Note. In assessment of young children (e.g., K-2), there are confounding developmental factors that affect the type of information best obtained and the most appropriate assessment methods.

• Instruction. Assessment for instructional purposes may include 1) decisions about teaching strategies and how well students are learning the specific learning targets set by the teacher or 2) specific individual diagnostic information for individual students. Decisions about classroom instruction are likely to be based on assessment information collected for or aggregated across the class as a whole or for groups of students and might include questions like, "Are my teaching strategies working?" or "Which of these two strategies is working better?" Diagnostic information about individual student learning involves collection of various pieces of assessment for a given student and might include questions such as "What misunderstandings does this student have?" Means of assessment for either classroom purpose can include both formal and informal strategies.

Assessments for instruction can help make decisions about teaching strategies, monitor how well students are learning, and give diagnostic information.

• Placement in Special Programs. A frequent purpose of assessment in schools is to qualify students for special programs. These purposes usually require specific types of standardized assessments, such as standardized individual achievement tests and intelligence tests. Although instructional recommendations often are made from these assessments, these types of assessments typically are not designed specifically for classroom instruction but rather to qualify students for programs such as exceptional children. This purpose may not align as well to the model described here, as conditions other than specific learning goals (e.g., handicapping conditions) may drive the assessments.

The ultimate point is ***no single type of assessment can meet all of the purposes of assessment or information needs of the various educational decision makers.***

An Environment for Classroom Assessment

Since different assessment purposes require different kinds of assessment and different assessment plans, assessment systems need to be developed that extend and expand beyond year-end state tests. Without question, day-to-day classroom assessment offers a powerful tool for the improvement of student achievement.

Creating classroom conditions where assessments and assessment results are used in positive ways, according to

Stiggins 1997 includes important underlying assumptions that help both teachers and students keep the focus on positive, constructive teaching-learning-assessment cycles.

• Assessments require clear thinking and effective communication. While many assessments do translate into scores, numbers are not the only way to communicate about achievement. Teachers can use words, pictures, illustrations and examples to communicate with students.

Further, strategies for communicating about student achievement are only as meaningful and useful as the students' understandings of what they are trying to accomplish and the quality of the assessments used. Teachers and students should be able to clearly respond to the questions, "What is being assessed here?" and "What do the results mean?"

• Classroom assessments have great influence on learning. On a day-to-day basis, teachers direct the assessments that determine what students learn and, to a large extent, how students feel about learning. Statewide accountability testing in North Carolina, as in most states, assesses students once a year at the end of the year. In contrast, classroom assessments are carried out by teachers in their classrooms from the beginning of the school year to the end. Further, classroom assessments are the ones that most likely "count" to students themselves. Students receive frequent feedback from teachers that have direct impact on their learning, motivation, and progress throughout the year.

Classroom assessments are most closely aligned with day-to-day instruction and are most influential in terms of decision making. Without question, teachers are the drivers of the assessment systems that determine the effectiveness of the schooling process.

• Students are key assessment users. There are many important assessment users at all levels of the educational process. However, only students can use the assessment results to set expectations for themselves. They decide how high to aim based on their sense of the likelihood they will succeed. They estimate their future success based on their prior classroom experiences (including grades). If those results reflect lack of academic success, teachers must act to change their

instructional approach to prevent the pattern of failure from becoming chronic. Ongoing classroom assessment and feedback offers a different way to reveal progress, success and potential to students.

• Clear and appropriate learning targets are essential. Teachers must first and foremost be confident, competent masters of the content they are teaching. Without a sense of the final destination reflected by clear teachers' standards, and sign posts along the way against which to check student progress, the teacher will have difficulty being effective. The quality of any assessment depends first on the clarity and appropriateness of the learning target.

Being a master of content empowers teachers to set signposts and benchmarks for learning.

• High quality assessment is a must. Unsound assessments can lead to misdiagnosed needs; failure to provide needed instructional support; use of inappropriate instructional approaches; counterproductive grouping of students; and misinformation provided to students, parents, and other decision makers. The purpose of this document is to present a picture of what high quality assessment looks like in the classroom.

• Assessment has personal implications for teachers and students. Teacher-student relationships have direct bearing on everything that happens in a classroom. Teaching and learning are very personal activities, if they are to be effective. Teachers must remain as objective as possible in the process of teaching and assessing.

Assessment exposes students to the possibility of academic and personal benefit and harm. In the face of assessment and evaluation, both teachers and students are vulnerable. Sound assessments link students to their constantly emerging sense of academic and personal self-efficacy. When they do not understand what is expected and how to succeed, students feel doomed to fail, and often give up.

• Assessment relates to both teaching and learning. Sometimes assessment is needed as a status check, to see where the student is in relation to the desired learning goal. At other times, assessment events can be powerful intervention tools. One of the most effective ways to do this is to bring students into the assessment process. Students who participate in the

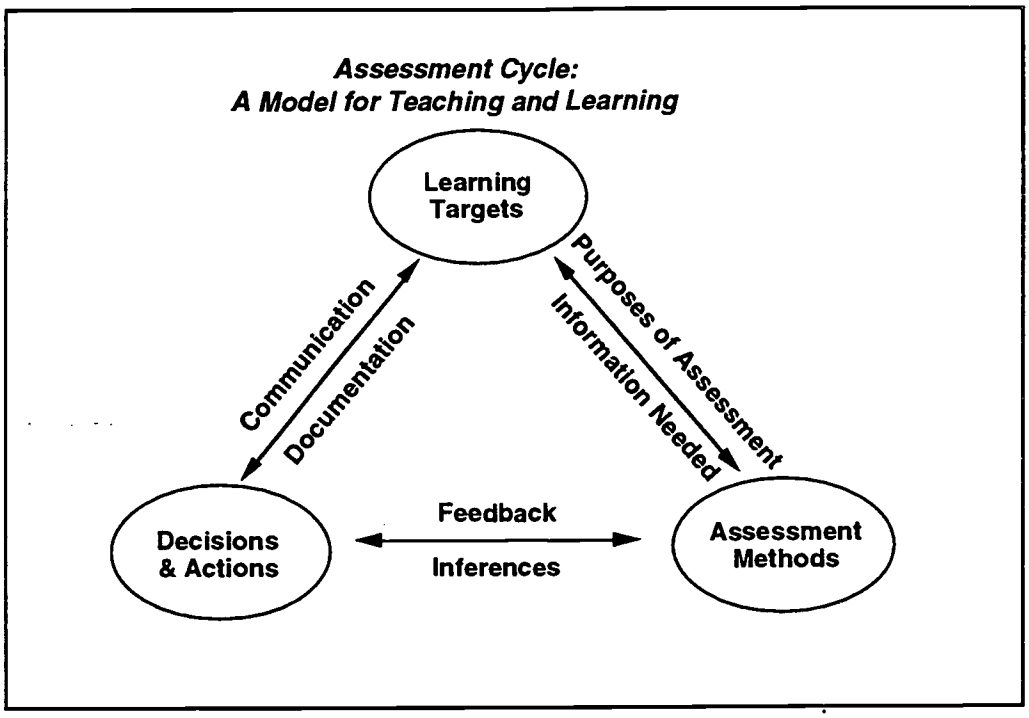
thoughtful analysis of quality work in order to identify and understand its critical elements become better at demonstrating their achievement and learning. They learn to identify and analyze their own shortcomings, take responsibility for improving them, and gauge their progress as they move forward.

The Assessment Cycle: A Model for Teaching and Learning

This model of assessment illustrates the essential components of the cycle of instruction and assessment and provides the foundation for this document. The key aspects of the assessment process appear in each of the three ovals in the diagram on the next page. The actions shown along the connecting arrows link the components. A brief description of each component is provided here, along with an overview of how it plays out for different purposes of assessment and at different levels of education. Sections that follow address more specifically the classroom assessment applications.

Learning Targets. The instructional purposes (i.e., the learning targets) should drive all assessment decisions. As previously noted, the nature of learning targets will differ greatly depending on the **purposes** of the assessment and the **types of information needed**. Assessments must be designed to match the learning questions of the user or they cannot provide useful information to educators. Although the nature of the targets will depend heavily on the specific objectives being taught, there are generic types of learning targets that cross subject matter. For example, many learning targets deal with *content knowledge*, either factual or conceptual. *Skills and processes* are important for application of knowledge. Still, students must engage in reasoning and problem solving using their content and skills. Other learning targets might involve student creations or *work products*. Finally, student attitudes are important learning targets, including such traits as persistence, interest in learning and the subject area, and confidence.

Each of these types of learning targets may best be addressed by different assessment strategies. However, it is often the case that assessments for accountability purposes will focus



primarily on learning targets for knowledge and/or skills and processes, and to some extent, reasoning and problem solving. As we move closer to the purposes of classroom instruction and diagnosis, learning targets are more numerous and more likely to include reasoning, products, and attitudes.

Learning targets are the goals and objectives in the *Standard Course of Study*.

It is important to define clearly our learning and achievement expectations for students. Students, as well as teachers, must have a clear idea of what they are expected to master and what success looks like. Assessment formats should reflect these expectations. Chapter 2 addresses learning targets in more detail.

Assessment Methods. Assessment methods should be selected to match the learning target. Methods of assessment are diverse and may vary by degree of structure (formal versus informal), method of demonstration (written versus demonstration or performance), and the extent to which they take the form of the learning target itself (direct versus indirect). In the classroom, the closer the form of assessment resembles the learning target, the more accurate the information will be for continued instruction. However, these methods often take more time, where other strategies may be less direct but more efficient. General categories of assessment strategies are identified here.

Assessments must be matched to the purpose and to the type of information needed.

- *Selected response* : Also referred to as "closed-ended" assessments. Includes multiple choice , true/false, matching items.

- *Written response*: Includes short-answer and essay tests. May include more extended written work, such as journals.

- *Performance*: Includes student demonstrations of learning, such as oral presentations, projects, dramatic depiction.

- *Conversations*: Includes student explanations of a task or teacher probe of a student's thinking to better identify how the student is understanding the learning targets; may be informal or specifically structured for diagnosis.

- *Observations*: Includes a wide range of opportunities for teachers to observe students' written, oral, and otherwise demonstrated work. Typically refers to carefully crafted observations to determine the extent to which students are mastering and understanding learning targets.

Referring back to the definition of assessment, the method(s) of assessment chosen should provide enough information for the user to make adequate inferences about student learning of the target(s). A quality assessment should provide a representative sample of student performances sufficient in scope to permit confident conclusions about student achievement. Several issues impact the defining of "a sufficient sample of performance" and the extent to which one can make confident inferences about student learning.

If an assessment is a less direct measure of learning (e.g., multiple choice test), several supplementary performance samples (e.g., observations, performances) may be required to conclude a student has mastered an achievement domain. The more closely an assessment method matches the learning target (i.e., the more the assessment directly measures learning), the less "inference" the teacher has to make in order to judge how well the target is learned.

Different assessment methods have different degrees of

reliability and validity for specific purposes. Thus, the kinds of inferences about learning may differ across methods of assessment. For example, the EOG multiple choice tests are very reliable for communicating about the learning of groups of students; but any single multiple choice test is less reliable for high stakes decisions about individual students, and it will not provide complete information on student mastery of an achievement domain. However, a teacher can use results of the EOG tests for her classroom to get a picture of how her students compare to other students in the same grade in the district or state. And it is one measure of individual student learning that can be combined with other samples of performance.

The higher the stakes of an action are for the student, the more reliable, valid, numerous, and diverse the assessment methods need to be for making decisions and taking action. For example, a promotion/retention decision should be based on numerous samples of student performance, including formal assessments and classroom evidence of student learning. On the other hand, inferences about how well students have learned an instructional unit may be based on observations of students' work and conversations about their understanding.

For on-going instruction, students and teachers need quality feedback about student performance.

Making Decisions and Taking Action. The ultimate reason for assessing students is to make various types of decisions that relate back to instructional goals and assessment purposes. The decisions made and actions taken can vary from knowing that an individual student or group of students has learned specific instructional objectives, to assigning grades to student work, to rewarding schools or posing sanctions for exemplary or poor overall student performance.

The purposes of the assessments determine how results of assessments need to be **documented and communicated**. For accountability purposes, communication to individual students may be less important or relevant to on-going instruction. While we *get feedback* about student performance from various assessments used, we should also give students feedback about their learning. Thus, teachers especially are called upon to communicate results of assessments to students, parents, and the school.

Documentation may be informal or more formal. Some forms

Use of technology can be extremely useful in gathering and documenting assessment information.

of documentation and communication have instructional purposes as well. *Portfolio assessment* is really a way of collecting and documenting student learning. The purpose of the portfolio (e.g., "best work" or learning progress) determines the types of information to be collected. While this is a means of documenting student learning, it provides evidence of learning, feedback to students and parents and a basis for teacher grading. Similarly, the newer practice of student-led conferences is a way of communicating results of learning that may also have instructional benefits. Students may learn to take responsibility for their own learning and to self-assess. As we devote more attention to quality assessment and systematic documenting, use of technology will be increasingly important.

Applications of the Model

The model for teaching and learning can apply to state, school, or classroom level.

State Level Application: Accountability

Learning Target: Students will read for the acquisition, interpretation, and application of information
Purpose/Information Needed: Accountability/Broad measure of learning
Assessment Method: Multiple choice test
Feedback/Inference: Overall level of achievement for schools
Decisions/Actions: Incentive money granted; low performing sanctions
Documenting/Communicating: Number and scores; designations of status

School Level: Program Evaluation

Learning Target: Students will read for the acquisition, interpretation, and application of information
Purpose/Information Needed: Determine how well students are mastering overall *Standard Course of Study* /EOG Test scores by grade level
Assessment Method: Multiple choice test
Feedback/Inference: Growth in grades for each subject area; how well each grade level in the school is succeeding in

promoting student mastery in both subject areas; areas of strengths and weaknesses
Decision/Actions: Does curriculum need to be realigned? Do we need professional development in particular areas?
Documenting/Communicating: Meetings with teachers explaining results; tables and graphs showing status by subject and grade level for further diagnostic work by staff.

Classroom Level: Instruction

Learning Target: Student will read for the acquisition, interpretation, and application of information
Purpose/Information Needed: Diagnose each student's achievement and determine appropriate sequence of instructional activities
Assessment Method: Multiple-choice and short-answer tests, supplemented with individual "Think Alouds"
Feedback/Inference: Strengths and weaknesses of student's comprehension in reading
Decisions/Actions: Comprehension of individual students; how to group students and sequence instructional activities
Documenting/Communicating: Conversations with students and parents; progress of growth documented with reading logs, Think Alouds, and other classroom assessments in literacy portfolios.

In a "Think Aloud" the teacher asks the student to read a passage and verbalize his or her thinking.

Final Note: Principles for Quality Classroom Assessment

Assessment literacy involves knowledge and skill. It means knowing how to gather dependable and accurate information about student achievement. It also means knowing how to use the assessment process and results to promote maximum achievement.

Quality assessments are a must in order to have dependable information about student achievement. How the assessment is conducted and how the results are used and communicated are also of crucial importance. Competence in the development of quality assessments and in the accurate use of assessment results is built around the principles that follow on the next pages. These principles are integrated into the Assessment Cycle Model presented in this document.

Principle 1. Quality assessments arise from and accurately reflect clearly specified, appropriate, and essential learning targets.

All educators must be competent in the content of the discipline that they are teaching. Knowing precisely what we are asking students to master is important because different targets require the application of different assessment methods. In any assessment context, educators must begin the assessment development process by defining a clear vision of what it means to succeed.

Principle 2. Quality assessments are specifically designed and focused to serve instructional purposes.

Assessments should not be designed without asking who will use the results and how. In education, there are generally three levels of use in terms of assessment results: classroom, school/ district, and state. Each important user of assessment information needs different information (often in different forms), at different times, to answer different questions. There is no single assessment capable of meeting the needs of all of the different levels of users.

To provide quality information for teacher, student and parent at the classroom level, teachers need sound classroom assessments that provide a variety of evidences of learning. To provide useful information at the levels of policy or program evaluation, educators may need quality standardized tests. Because of the differences in information needs, educators must begin each assessment event with a clear sense of whose needs are being met. Otherwise, assessments may not fulfill their purpose.

Principle 3. Quality assessments accurately reflect the intended target and serve the intended purpose.

Since there are different kinds of learning targets and no single assessment method will reflect them all, teachers must rely on a variety of methods. The options include selected response (multiple choice, true/false, matching, and fill-in the blank), written responses, performance assessments or demonstrations (based on observation or judgment), and direct personal communication with the student. The assessment challenge is

to match the most appropriate method with an intended target. The professional development challenge is to be certain that all concerned with quality assessment know and understand how the various pieces of the puzzle fit together.

Principle 4. Quality assessments promote equity by providing ample opportunities for students to demonstrate their learning.

Students bring to the classroom varied and unique experiences. Equitable assessments focus attention on each student's learning as the teacher simultaneously assesses the class as a whole. Expectations for achievement as demonstrated through high quality performances remain constant for all students. Modifications of assessment strategies or format, however, honor students' special needs. Just as opportunities to learn influence the quality of education of any individual student, opportunities to demonstrate learning are likely to influence the actions and decisions that educators make related to individuals.

Principle 5. Quality assessments provide sufficient evidence of learning to permit confident conclusions (inferences) about student achievement.

Classroom assessment should accurately represent the body of knowledge and applications the students were expected to learn. If assessments do not accurately reflect what was taught, the assessment will yield inaccurate information about the student and may misrepresent that student's real learning.

The question must always be posed, "Have I gathered enough information of the right kind, over a long enough period of time, to draw confident conclusions about the student's learning?" Teachers must adjust their sampling strategies so the assessment context varies to produce maximum quality results at minimum cost and effort (for both the teacher and the student).

Principle 6. Quality assessments are designed, developed and used in ways that eliminate bias or distortion that may interfere with the accuracy of results.

Even if all other quality standards are met, problems can still

arise from the test, the student or the environment in which the test is administered. For example, tests can consist of poorly worded questions, place reading or writing demands on respondents that are confounded with mastery of the material being tested, have more than one correct response, be incorrectly scored, or contain racial, cultural, socio-economic, or gender bias. The student can experience extreme anxiety or interpret test items differently from what was intended, or can cheat, guess, or lack motivation. Further, the assessment environment could be uncomfortable, poorly lighted, noisy or otherwise distracting. Any of these variables could give rise to inaccurate test results.

These core principles, discussed as standards for quality assessment by both Stiggins (1997) and *Assessment Standards for School Mathematics* (National Council of Teachers of Mathematics 1995), can help direct development of appropriate classroom assessments by teachers and other educators.

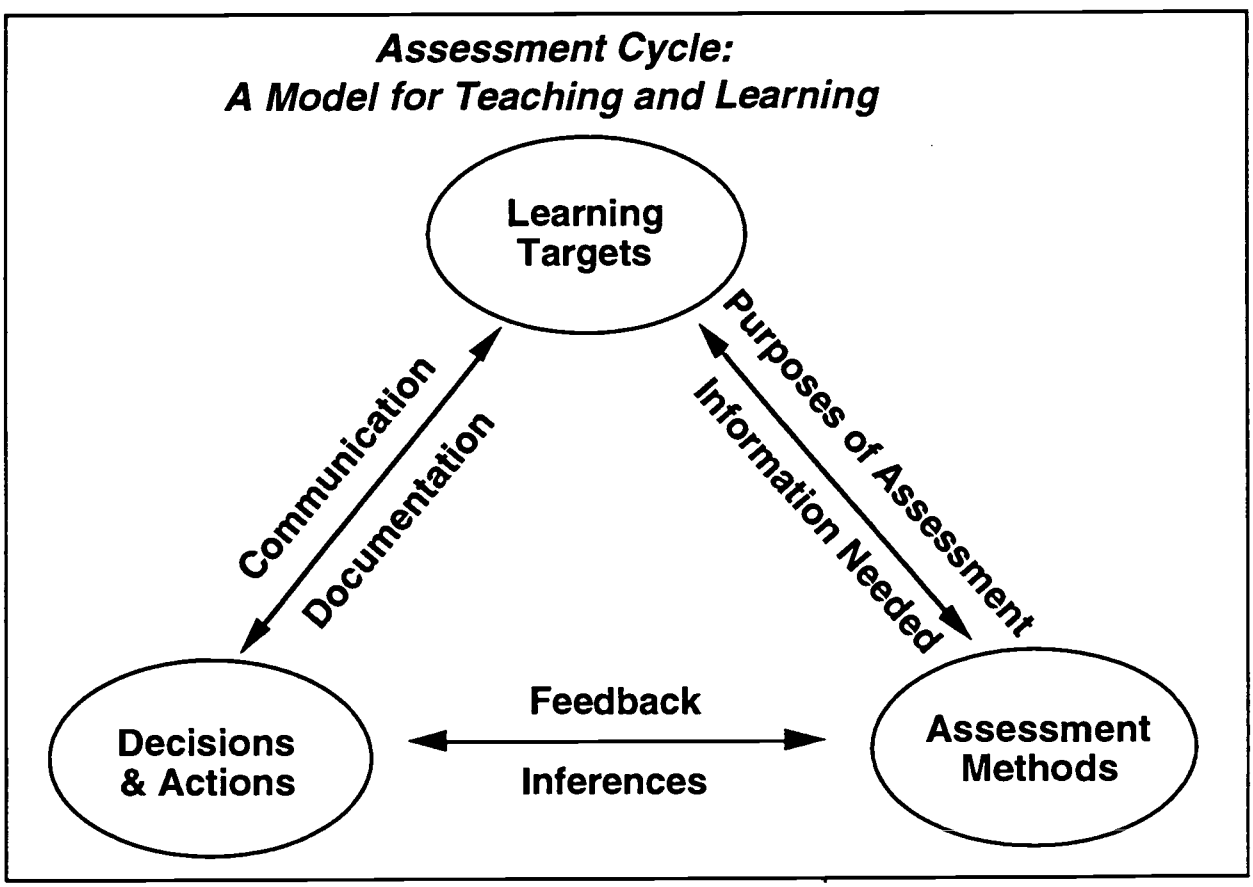
Chapter 1

Classroom Assessment and Instruction

What is classroom assessment?

Classroom assessment involves teachers and students in an on-going process of gathering information to make inferences and decisions related to instruction and student learning. A strong model for teaching and learning involves the following cycle of classroom assessment.

"Assessment" comes from a root word meaning "to sit beside."



Learning Targets

In the process of classroom assessment, effective teachers possess a conceptual, in-depth understanding of the content they are teaching. They are able to identify what learning targets are most important. They establish priorities and categories of targets taken from the *Standard Course of Study*. They decide how to break down those categories into clusters of manageable, appropriate instructional units. Most importantly, they clarify for themselves, and for their students, what those learning targets are and what mastery of them will look like.

Having a thorough, deep understanding of the content we are teaching is critical. Because the content knowledge often changes, we must be continuous learners to be good teachers.

Clarifying goals (learning targets) and planning for instruction are not simple tasks. These actions involve knowledge of what students need to learn, how students learn, and the individual interests and work habits of the students within the classroom.

Teachers plan broad, yearlong goals as well as how they will teach and sequence components of these goals throughout the year. For example, in mathematics, a broad goal might be the development of place value understandings. Teachers attend to skills, such as meaningful counting using one-to-one correspondence. As they count, students learn that a number can name the last object counted in a group but will also include the entire group of objects. Along the way students learn that counting the same group of objects by ones or two's or five's will give the same answer to the question "how many." They learn that ten ones and one ten are equivalent and that the position of a digit determines its value.

Clear learning targets are essential for classroom assessment.

In reading, a broad goal may be to use different reading strategies when reading for different purposes. For example, looking at subtitles, pictures and charts to clarify meaning may work well in reading a textbook. Visualizing what the scene or the main character looks like may work well in reading a novel or short story. Teachers make sure that students are learning about and practicing relevant reading strategies in the context of what they are reading.

Teachers incorporate many different types of learning targets into their lessons. At times the goal is for students to learn a new concept or to gain specific content knowledge. Other lessons focus on the acquisition of skills. Many lessons are designed to develop students' reasoning and problem-solving

abilities. Still other lessons are planned to help students learn to apply their knowledge and skills in order to create worthwhile products. All of these content-based lessons incorporate goals related to students' work habits. They also include goals related to positive attitudes toward learning the specific content (e.g., mathematics) and building self-confidence.

Purposes and Information Needed

The primary purpose of classroom assessment is to direct and inform student learning.

Within the classroom, teachers and students are the primary users of assessment information. The most important purpose is to direct and inform student learning. The teacher and students need lots of information about each student's understandings and performances. They need more than one "snapshot" or one piece of data; they need multiple evidences of learning that will allow diagnosing, monitoring progress, evaluating achievement, and deciding about future instruction—for each student!

The various purposes of classroom assessment, appropriate for all content areas and all grades, are present throughout the school year. They relate to the class as a whole and to individual students:

- Diagnostic assessments provide teachers with information about what students already know;
- Formal and informal assessments help teachers know what to plan next for instruction;
- Assessments monitor the individual student's progress toward the specified goals;
- A variety of assessments helps teachers evaluate learning and assign grades;
- Summative assessments measure achievement before teachers move to the next topic.

The essence of assessment is constructive feedback.

Classroom assessments are formative when they are used during instruction to monitor student growth and progress and summative when they are used at the end of instruction to evaluate student achievement. When assessment is summative, teachers "put on a different hat" and judge (evaluate) the students' work in relation to criteria (benchmarks). Feedback frequently takes the form of a grade.

These assessment purposes are not mutually exclusive; the

various assessments administered in the classroom all serve to give the teacher information about what students are thinking, what they can explain, and what they can apply. The same assessments allow students to self-assess and examine their progress toward learning targets.

Teachers create conditions that produce confident and competent learners through the opportunities they provide for students.

Important, too, are the goals students set for themselves and their own **self-assessments**. Teachers should always ask students to assess themselves. Later in this document, there is more discussion about self-assessment, including creating and using different methods of assessment, keeping assessment records, and creating scoring guides (rubrics).

> Take Time to Reflect
>
> - Do I consider various purposes when I am planning assessments?
> - Do I provide opportunities for my students to assess themselves?

Assessment Methods

Assessment methods should be on-going, purposeful, and diverse.

As part of this cycle teachers also use a variety of formal and informal assessment methods to gather evidence of student learning. They match the type of assessment method to the learning target they want to measure.

Thus, in planning for assessment, teachers need to have knowledge about different forms of assessment. Exploring one new assessment method at a time seems to be a reasonable approach to learning new techniques. Teachers can introduce "controlled variety" of assessments by:

- Choosing a new assessment strategy carefully,
- Implementing it systemically and thoughtfully,
- Reflectively analyzing its effectiveness.

Becoming comfortable with an assessment method means that, through study and experience, teachers find out what kind of information can be efficiently gathered through that method, the time involved in using the method, how easy it is to use, how easy it is to score, and how easy it is to interpret the results. Several types of methods are identified on the next page.

Selected response includes:
- Multiple-choice,
- Matching,
- True-false questions.

Written response includes:
- Students' drawings,
- Open-ended questions,
- Journals and learning logs,
- Short answer,
- Discussion questions.

Performances include:
- Performance tasks,
- Oral presentations,
- Projects.

Conversations (interviews) occur in:
- Whole class settings,
- Small groups,
- Individual interviews.

Observations can include:
- Formal, prompted behaviors,
- Informal, unprompted behaviors.

In our discussion, assessment strategies are the five large categories. Assessment methods are the specifics under each category.

Teachers use a variety of formal and informal assessment methods to gather evidence of student learning in a variety of contexts, over an extended period of time.

Take Time to Reflect

- Which of the assessment strategies listed above would I like to incorporate into the strategies I am currently using?
- How will I implement it systematically and thoughtfully?
- How can I analyze its effectiveness with students?

Feedback and Inferences

The essence of assessment is feedback—to the teacher, to the student and to other interested individuals like parents. The value of the assessment is thus in the **relevance**, the **precision**,

and the **accuracy** of the information that is extracted.

As they gather the evidence, teachers ask reflective questions of themselves. For example, they may ask:

Reflective questions help teachers to get good information in order to make appropriate inferences.

- What do these errors actually tell me about the students' thinking and understanding?
- Do I have sufficient evidence to know how well the students really understand?
- How well can I generalize about how much students can do?
- What other evidence may I need?

Reflective questions such as these lead the teacher to decide how to make sense of the evidence of student learning. They allow the teacher to decide **what information and feedback can be extracted from student assessment data and what inferences and interpretations can be made about student learning.**

Imagine these scenarios: Classroom instruction and assessment began by clarifying specifically what was to be learned. Decisions then were made for how to best gather information that reveals what students know and can do and what they still need to learn. Throughout instructional and assessment activities, data have been collected. Teachers now must decide what is important about the information they have. To make those decisions, different teachers may find that they are asking themselves different questions.

Ms. Parker: "From the students' responses, what do I know about the students' understandings? How complete are the important ideas the students already have and how am I going to build upon them?"

Mr. Bradbury: "What does it mean when a student answers factual questions about a population graph but is unable to make summarizing statements or reasonable predictions?"

Ms. Martinez: "In solving problems, do the students solve one type of problem but consistently miss another?"

Mr. Wu: "Is the student able to describe a passage in very literal terms but unable to recognize a story as an allegory? What does this tell me about the student's reading

comprehension? About the student's understanding of allegory?

After information has been extracted, teachers are ready to make broader interpretations about the results. Teachers will interpret assessment information as it relates to teaching the class as a whole and to working with individual students. This important aspect of the instructional cycle calls on teachers to be reflective as they condense multiple forms of assessment results collected over time. Seeing important patterns becomes crucial, as does the consideration of formal as well as informal assessment. For example, the verbal and nonverbal clues students give related to their thinking, the nature of errors that students consistently make, and the misconceptions–not just the incorrect answers–all help to inform teachers about student learning.

Classroom assessment is not an add-on activity; many times it is indistinguishable from instruction.

Feedback to Students

Providing students with feedback about their progress toward the learning targets is another action that arises out of using the results of assessment. Helping students become active, responsible partners in the assessment of their learning is likely to increase student achievement because students are able to see their own progress and set intermediate goals for themselves.

Feedback to students takes many forms. It can be through scheduled conferences with students or informal conversations. It can be written on individual's papers or given to the class as a whole. Feedback should often take the form of questions designed to prompt students to make decisions rather than giving them directives about what to do. Questions encourage students' thinking about what they know and are doing so that they recognize inconsistencies or mistakes.

When students are first learning new content, usually they are not completely successful and need specific guidance to help them know what they need to do to get better. Telling students to "try harder" or that they did a "good job" does not give them information that is useful for the next assignment. Likewise, a numerical score or letter grade does not let students know what to do differently (what they should continue or what they need to change).

Praise is encouraging to everyone and should not be dismissed. However, pointing out strengths of a performance- or better still, helping students to self-assess and identify the strengths - moves them toward competence with confidence. Models to compare their work with and/or rubrics are sources of information that give students ways of looking at their own work and deciding how to make it better.

Time used for discussions with students about their progress toward learning targets is time invested wisely.

Monitoring students' progress involves sharing with families. Students and families both need to know about the progress the student is making toward attaining the learning target and recognize accomplishment of benchmarks along the way. Monitoring also raises "red flags" if students are not making appropriate progress. Assessment data help teachers and families decide if students need specific intervention beyond the regular classroom and what that intervention should look like. These data also help to identify special talents.

There is no doubt that the decisions teachers make from assessment data are as critical as the need for quality information itself. Knowing when to intervene and when to wait and see is difficult. Too much intervention may result in students only memorizing and parroting back what the teacher has modeled or said. Too little intervention may result in students "practicing it wrong." Both wrong actions and inaction may result when classroom assessment is not wisely designed and used. Feedback to students, beyond just assigning a grade to work, is an important component of learning and should receive increasing attention.

Ineffective Feedback:	**Effective Feedback:**
Not Clear: "Try harder."	Clear: "You left out two sections."
Not Precise: "Your work is improving."	Precise: "You made a left turn instead of a right turn."
Not Related to Criteria: "You're better at this than most students."	Related to Criteria: "Your summary was organized--especially in these places."

Take Time to Reflect

- Is my feedback to students effective or ineffective?
- How can I make my feedback to students more effective?

Making Decisions and Taking Action

The next step in the classroom assessment process is to decide upon a course of action. Because assessment is used for different purposes within the classroom, the results of assessments are used in different ways.

There are times when teachers decide to "file away" the information they have gathered because it confirms plans already made or it is inconclusive. Teachers may plan whole-class lessons or group students for specific instruction. They may decide to work with individuals. The positive result of classroom assessment is that **decisions and actions are based on knowledge of student progress in relation to the learning targets.**

For example, the teacher may decide to reteach key concepts, to move to the next unit of instruction, to regroup students for further instruction, or to allow more practice and application time. The cycle may then begin again, with the teacher's identifying and clarifying the next learning targets.

Communication/Documentation

However, before the cycle of classroom assessment begins again, the teacher needs to document the students' learning and communicate to students. Documentation of student learning can include checklists, anecdotal records, observations, grades, and portfolios. At the end of each grading period information about student learning is synthesized and communicated to students and parents on report cards. All communications should provide clear, precise, usable information to students, parents, administrators, or other

interested adults. Communication can be informal (a telephone conversation, note or informal conference) as well as formal (a report card or scheduled conference).

Preparation, Reflection, and Decisions

Classroom assessment involves *preparation*. Throughout the year, teachers continue to clarify targets from the *Standard Course of Study* before planning instruction, before deciding the purpose and the most appropriate type of assessment for the instructional goal, and before planning how best to do those assessments. Classroom assessment also involves *reflection*. In every phase of the process, teachers examine what they are learning about students' thinking and reasoning and the students' abilities to apply their knowledge and skills. Finally, classroom assessment involves *decisions*. As teachers and students give meaning to what they have learned, they take action: both teachers and students ask more questions, gather further evidence, and give and receive feedback. Teachers plan further experiences, evaluate the assessments, and, at appropriate times, assign grades.

Classroom assessment is what effective teachers have always done.

Classroom instruction can look very different in different contexts. Teachers have different teaching styles, classes of students have different instructional needs, and teachers at different levels have different goals. The vignette on the next page, " Research Assignment: Two Case Studies" illustrates how different teachers use the assessment cycle in different ways.

A Research Assignment: Two Case Studies

Ms. Jones has learning targets for her class that are clear and well articulated. She wants her students to be able to use multiple sources to find information, prioritize that information into concepts of greater and lesser importance, find details that substantiate concepts, and organize and write their ideas clearly. In assigning her students a research project, she gives specific assignments and due dates, as well as precise directions on what she expects from her students.

Ms. Jones sets requirements for the number and categories of references they must have (books, periodicals, sources from the Internet, etc.). She teaches students to use note cards and sets a due date for when a specific number of cards are due and the format they must have. She requires a formal outline, rough drafts, and a final copy. Students know clearly what she expects and the timeline for producing it. They also know how they will be assessed—Ms. Jones shows students models of excellence (student outlines, final copies, etc.) and grades all assignments with a rubric and criteria, which she gives to the students at the beginning of the unit. Ms. Jones is thus very directive in her instructional unit and her use of the classroom assessment cycle.

Ms. Smith has the same learning targets as Ms. Jones. However, Ms. Smith is much less directive with her students. She begins by showing students three research papers that were written by former students—they range in quality from excellent to needing a great deal of work. All students read all three models, discuss their quality, and brainstorm a list of characteristics of a good research paper. In this way, Ms. Smith makes sure that students begin to conceptualize what the target looks like—what are the criteria for excellence. From class discussion, the students generate the rubric they will use for evaluation the research paper and the research process.

During the unit, Ms. Smith allows students a great deal of flexibility. They may use any system of taking notes as long as they can articulate what they are doing and how well it is working. Ms. Smith allows students to choose how to organize their ideas. They do not have to create a formal outline; they may use a web or graphic organizer as long as they can discuss the organizational pattern they are using. She requires that students rewrite a rough draft, even several times, if the drafts do not meet the minimum standards according to the class rubric. Ms. Smith also asks students to self assess and peer assess the process and products of research.

Although these two teachers are quite different, their approach to the teaching-assessment cycle has many commonalities:

- The learning targets are worthy of time and effort, not trivial or narrow.
- The targets are clear and understandable to students.
- There are multiple assessments during the unit—not just one grade at the end.
- The feedback to the students is clear and useful.
- Teacher decisions are then based on accumulated assessment information.
- Both teachers have an organized system for documenting progress and communication with students.

What is different about classroom assessment today?

Classroom assessment is what effective teachers have always done. They plan. They decide where and how to begin. They teach. They try to figure out what the students are thinking and if they are learning the intended targets. They make decisions about whether to give students more time and experiences or to move on. Periodically they evaluate student progress, making judgments that result in grades, reports, or special placements.

What is different today in classroom assessment is a change in emphasis:

- Learning targets are clearly described in advance and accompanied by equally clear performance expectations and public criteria.
- In addition to correct answers, there is a focus on students' thinking and the processes used by students in arriving at their answers.
- Evidence of learning is gathered in multiple ways over time rather than making decisions based on single tests at one point in time.
- Students assume more responsibility for their own learning through self-assessment, portfolios, student-led conferences, and other formats.
- There is an expectation by the public for more consistency between teacher-assigned grades and external assessments of student learning, such as state tests.

Teachers can enrich their classroom assessment practices by entering the assessment cycle at any point. Teachers can begin by reflecting on what is currently happening compared to what they would like to happen. Supportive colleagues are also helpful, as implementation of changes is not easy.

Classroom Instruction and Assessment: A Continuous Cycle

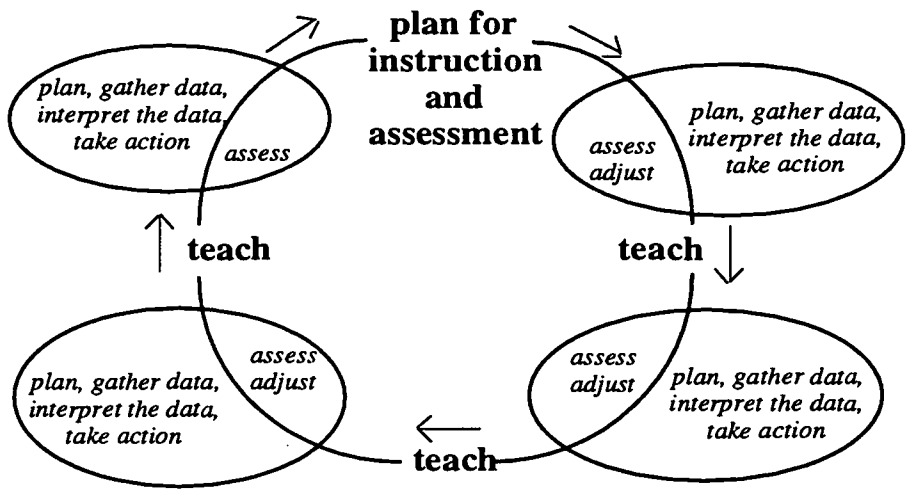
The diagram on the next page illustrates the recursive nature of instruction and assessment. Moving clockwise, one can see there are assessment cycles integrated into the instructional plan. All along teachers are observing, assessing, interpreting

what they have learned, and adjusting their teaching plans. At some point they assess summatively, noting where children are in the process of attaining the learning target so that future instruction will meet their needs.

Classroom assessment is not an add-on activity; many times it is indistinguishable from instruction. This especially is true if plans for lessons and plans for assessment are made together. Teachers should know why they are assessing and modify the ways they will gather information accordingly.

Thus, at the beginning of each cycle of instruction there may be a mini-cycle of diagnosis. At times teachers may carry out this diagnostic assessment very informally, looking over work from related units and drawing on classroom observations. Other times a written assessment or structured task is an important part of the planning process.

Classroom Instruction with Assessment in the Continuous Cycle



A Blueprint for Classroom Assessment

Stories and questions often help us think about what we are doing in different ways. The following questions may guide our thinking about classroom assessment.

- Why are we assessing?
- What are we assessing?

- When do we assess?
- Where do we assess?
- Whom are we assessing?
- How are we assessing?

Consider the following vignette of two teachers discussing the questions behind a "Blueprint for Classroom Assessment":

Teacher A: " I 've changed my purposes for assessing students. I used to think about giving grades a lot. Now when I think about ***WHY*** *I assess, I realize it's not just for grading but to see how well students are progressing, to judge if we can go on or if I need to reteach, and to help students to understand and to monitor their own learning."*

Gathering information and finding answers to questions about students' learning is what classroom assessment is all about.

Teacher B: " True, but the biggest change I've made this year is to think more about **WHAT** I'm assessing. I don't just want to test for factual information. I want to give new problems to see if my students are learning how to reason. And I want to see if they are developing the attitudes they need and the ability to apply information."

Teacher A: "I don't rely as much on "tests" as I used to. While "tests" occur periodically in my classroom, now I think as much about ongoing assessments like student journals, my observations, conferences, and classroom discussions. Sometimes I'm diagnosing, sometimes I'm giving an interim assessment, and sometimes the assessment is for a final grade. But I've learned that ***WHEN*** *I assess is important—and my students and I try to assess progress every day."*

Teacher B: " I know what you mean. I used to think of a student conference or a classroom discussion as just instruction. But it's also a chance for me to assess student understanding and to help students self-assess. Assessment is taking place **WHEREVER** I interact with and observe my students."

Teacher A: "I think it's important to look at ***WHOM*** *I'm assessing—to consider my students first and foremost. I have to think about what their experiences have been, what their interests are, and what I already know and don't know about them. That helps me to plan my classroom assessments and to understand how to interpret standardized assessments."*

Teacher B: "Knowing that I have different purposes for assessment is one of the reasons I've changed **HOW** I assess. I used to give only pencil and paper tests; now I also use videotapes, portfolios, anecdotal records, discussions, and student-led conferences."

Another way to think about the recursive nature of classroom assessment is with the following analogy:

Classroom assessment is like a fractal. It is a pattern that is repeated over and over throughout the school year. Each individual cycle is a miniature of the larger piece, and each larger piece is made up of smaller versions of the whole.

fractal (frak'tōl) n.

A geometric pattern that is repeated at ever smaller scales to produce irregular shapes and surfaces that cannot be represented by classical geometry.

Fractals are used especially in computer modeling.

Examine a head of broccoli for an everyday example of a fractal.

During the school year, grade-level goals are identified, taught, and assessed at the end of the year. To accomplish the grade-level goals, specific learning targets are identified and organized into instructional units that exemplify smaller iterations of the process. Within these instructional units teachers clarify components of learning targets and create short-term goals and assessments to better plan for teaching, learning, and assessing.

Teachers' Voices: The Importance of Classroom Assessment

"Since the training in classroom assessment, I have used my *Standard Course of Study* as my guiding light. I'm not doing those cute things that aren't in the course of study just because they are cute. I am also paying closer attention to higher order thinking."

" Primarily I appreciate the emphasis on aligning mandated curriculum with what and how I teach and with how I make

students responsible for the material. I have reviewed several rubrics and project assignments used in the past trying to eliminate "fat." By "fat" I mean activities that in the past I may have had students do that did not have clear benefits as related to the curriculum."

These quotations from teachers reflect their feelings about classroom assessment.

"My focus has begun to shift from seeing the class as an instructional group to seeing it as a group of individuals with different capabilities that had to be assessed as such."

"The impact of the assessment changes I used in my classroom was most apparent in student attitude. The children enjoyed knowing exactly what was expected of them during an activity in the classroom. I found their behavior was more on task and that the work products produced were much higher quality products than I had seen previously from the same students. They also became much better thinkers and decision-makers as they took responsibility for the activities we completed."

"As students took a more active role in their learning. I saw attitudes and effort change. Students became interested in learning for the sake of learning. They began looking at the progress they had made rather than just "waiting" for the grade. They knew the state tests were there in the distance, but they didn't come to school to perform on those tests. They came to school because learning was self-motivated."

"Designing a quality assessment takes practice. It is a developed skill. It has been like learning to ride a bike; once you learn how and practice, you never forget."

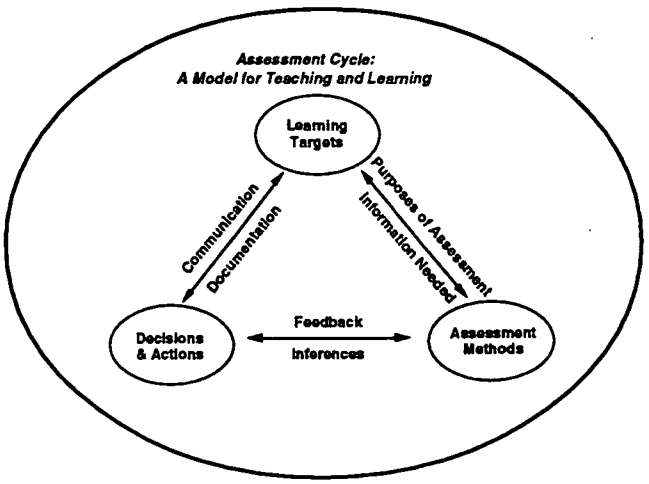
"It has taken me a long time to learn that I can assess while I also teach. But once I understood that idea, I found that assessment became so much easier and so much more meaningful"

"One impact I have seen this year is an increase in the use of higher order thinking skills in my classroom. I believe a combination of factors caused this. One factor is making clear the standard for student achievement. The other has been the use of on-going assessment."

"The impact on my students in changing my assessments, even in a limited way, has been greater than I ever imagined."

Chapter 2

Clarifying Learning Targets



"Identify your destination before you begin the journey."

It is extremely important for teachers to identify clearly the important learning targets for students before planning or beginning instruction. Teachers should ask themselves these questions *in the following sequence*:

1. What is it I want my students to understand and be able to do?
2. How will I know when they understand and can do those things?
3. What instructional activities will best teach my students these things?

This sequence may seem reversed—many teachers first decide what and how to teach and then what and how they should assess. However, as the fable on the next page shows, you need to identify your destination before you begin the journey of going there!

Beginning the journey:
- Identify learning targets
- Clarify what to assess
- Clarify how to assess
- Plan instruction to meet curriculum goals

Once there was a little seahorse who set out to seek his fortune. He swam through the ocean, proud of himself and the mission he was undertaking. Soon he met a clam.

"Psst, hey bud! Where are you going?" asked the clam.

"I'm going to seek my fortune," replied the little seahorse very proudly.

"For two pieces of gold," said the clam, "I will sell you these handy dandy flippers and you can zoom through the ocean twice as fast."

"Wow!" said the little seahorse as he paid the two pieces of gold. He put on the flippers and zoomed off into the ocean twice as fast. Soon he met a sponge.

"Psst! hey bud! Where are you going?" asked the sponge.

"I'm going to seek my fortune," replied the little seahorse very proudly.

"For ten pieces of gold," said the sponge, "I will sell you this handy dandy scooter and you can zoom through the ocean ten times as fast."

"Wow!" said the little seahorse as he paid the ten pieces of gold. He jumped on the scooter and zoomed off into the ocean ten times as fast. Soon he met a shark.

"Psst! hey bud! Where are you going?" asked the shark.

"I'm going to seek my fortune," replied the little seahorse very proudly.

Cunningly the shark said "I know the greatest short cut. I will carry you there and you can get there fifty times as fast." He opened his mouth and pointed inside.

"Wow!" said the little seahorse as he jumped on the handy dandy scooter and zoomed off into the shark's mouth —there to be devoured.

The moral of the story: if you don't know where you are headed, you might end up someplace else and not even realize it.

Retold from a story by R.F. Mager in *Preparing Instructional Objectives*, Palo Alto, CA:Fearon Publishers, 1962.

Using the *North Carolina Standard Course of Study* to Clarify Targets

In North Carolina teachers are responsible for teaching the content specified by the *Standard Course of Study*. In every discipline, the content builds from one grade to the next. The content may be examined vertically as a continuum which follows learning from kindergarten through the twelfth grade or horizontally by specific grade levels. Careful examination of the goals and objectives across grade levels shows the big ideas.

While identifying the content, the *Standard Course of Study* does not specify how the content standards, goals, and objectives are to be taught. The teacher decides the specific

activities, when objectives should be addressed, how objectives will be grouped for instruction, and the amount of time devoted to different objectives. In this way, teachers are able to factor in their knowledge of the students' interests and backgrounds as they make plans for instructional units throughout the year.

The overall goal is that **by the end of the year**, all students will have become proficient with the content described for their grade level. "Proficiency" means that they can model and explain the concepts, they can use the content appropriately and accurately, and they are fluent and comfortable in applying the skills related to the discipline.

All assessment should have positive consequences for students

Knowing what is to be learned is the starting point. Teachers make general, long-term plans before school begins. That is, in order for students to accomplish the goals in the curriculum, teachers break the year's content into manageable chunks. They plan what will be taught in each grading period. Short-term goals are established as units of instruction begin. Day-to-day plans are created for each week. During lessons, teachers make moment-to-moment instructional decisions.

When we let our students see our vision clearly, we eliminate the mystery surrounding success.

Types of learning targets

Teachers recognize that there are different types of learning targets, some more long-term than others:

- Content knowledge, factual or conceptual,
- Reasoning and problem solving,
- Skills and processes,
- Ability to create products and make applications,
- Attitudes.

Teachers need to consider the differences in these kinds of learning targets before they prioritize what are the most important goals of the year.

Content knowledge targets, which are important in every content area, include more than facts; they also include concepts and generalizations. For example, in English Language Arts, students need to know factual information about how to use punctuation; they also need to understand the concept of adjusting their spoken and written language to communicate with different audiences for different purposes.

In mathematics, students need to know their times tables; they also need to understand generalizations about the importance of geometry in the modern world.

Reasoning and problem solving targets, which ask students to think about and apply knowledge and information, are also important in every content area. In fact, most disciplines have an accepted way of thinking in the content area, so we can talk about historical thinking, scientific reasoning, and mathematical problem-solving. It is important to realize that human beings do not reason in isolation—they reason about something *in a particular context.* Students need to learn how to reason in the context of the particular problem or topic and in the context of the content.

Skills and processes involve specific, definable behaviors that may be complex. For example, students may be asked to learn processes such as how to read fluently, to speak articulately, to work cooperatively as a team member, to use a microscope properly, or to solve an algebra equation. The only way to determine if students have actually mastered these skills or processes is to assess their performance.

Ability to create products and make applications is another target valued in most disciplines. For example, completing a research paper, designing a science project, producing art work, writing a paper, and compiling a portfolio all require that the students create a quality product.

Attitudes and disposition (work habits) often are overlooked because they are difficult to assess. However, they are extremely important to student success. Perseverance in problem solving, the willingness to be fair minded when encountering new ideas, curiosity, and intellectual integrity are necessary for student success in learning.

Take Time to Reflect

The chart on the following page is adapted from the work of Rick Stiggins.

- Identify specific examples of each category in your discipline.

Important Learning Targets

Category	Examples
Knowledge and information	vocabulary, concepts and big ideas, historical facts, symbols for chemicals/elements
Reasoning/problem solving	analyzing, comparing, inferring, evaluating
Skills/processes	reading strategies, speaking, interacting with others, motor skills, science lab procedures, applying mathematical formulas
Products and applications	writing samples with specific attributes to be created, research reports, videos, art products, woodshop products
Attitudes or disposition	curiosity, open-mindedness, perseverance, responsibility

Questions in the Classroom Instruction and Assessment Cycle

Within the classroom these learning targets are likely to overlap. For example, making a quality product will involve skills and processes, as well as sound reasoning. However, these categories may help the teacher to move beyond targets of knowledge and information and they may help to clarify answers to the essential questions introduced at the beginning of this chapter:

- What do I want my students to understand and be able to do?

- How will I know when they understand and can do those things?
- What instructional activities will best teach my students these things?

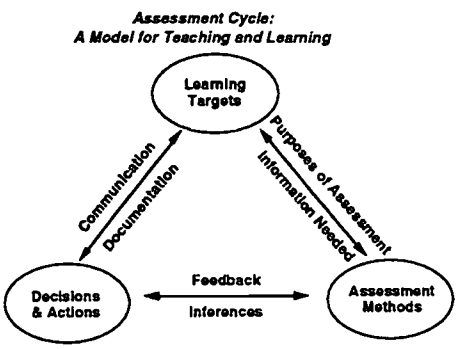
In the classroom instruction and assessment cycle, teachers find that, in addition to the three preceding essential questions, other guiding questions occur over and over:

- *What types of learning targets am I establishing (i.e., knowledge? reasoning? skills? etc.)*
- *Am I communicating clearly and precisely what I want my students to understand and be able to do?*
- *What are the intermediate steps along the way to becoming proficient with the content?*
- *What are the likely misconceptions students may have?*
- *How will I use the results of this assessment?*
- *How will my students use the results of this assessment?*

Establishing clear learning targets means that teachers will answer questions like these with specific examples and descriptors so that students and parents will know what accomplishment of the goals and objectives in the *Standard Course of Study* looks like. It means organizing the goals and objectives in ways that are comfortable and interesting for the teacher and the students.

Good instruction and assessment thus begin with clear learning targets for both the teacher and the student. This means that students can not only identify learning targets, they can also describe them and visualize what mastery of the target would look like.

Clarifying Learning Targets is a Crucial Step in the Classroom Instruction and Assessment Cycle



Take Time to Reflect

1. Think about the major goals that you want students to learn. Use the *North Carolina Standard Course of Study* to help you identify those major goals.

2. Describe what students should be able to do or explain when they have attained these important learning targets. What does it look like when students have achieved this goal?

3. If possible, select quality samples of student work that show what achieving this learning target looks like. What are the criteria of the work that achieves quality?

4. Can this goal be subdivided into smaller objectives? What does it look like when students have obtained the objectives?

5. Can you group these goals and objectives into categories? Can you cluster them into manageable, connected, logical units?

6. Is there an order of importance of the categories?

7. How are these objectives connected? How can you help students make connections among them?

8. Can you model, explain, illustrate and describe these goals and objectives so that your students can understand them clearly?

Levels of Thinking and Reasoning

In the *Standard Course of Study* the levels of thinking and reasoning are classified in the following way. Verbs which are likely to be used in questions and directions related to these classifications are included below.

Knowing: Recalling or remembering information

list, name, label, recall, identify, match, choose

Organizing: Arranging information so it can be used effectively

categorize, group, classify, compare, contrast

Applying: Using information for practical purposes

apply, make, show, record, construct, demonstrate, illustrate

Analyzing: Examining parts and relationships

outline, diagram, differentiate, analyze

Generating: Producing new information or ideas

conclude, predict, explain, elaborate, infer

Integrating: Connecting and combining ideas

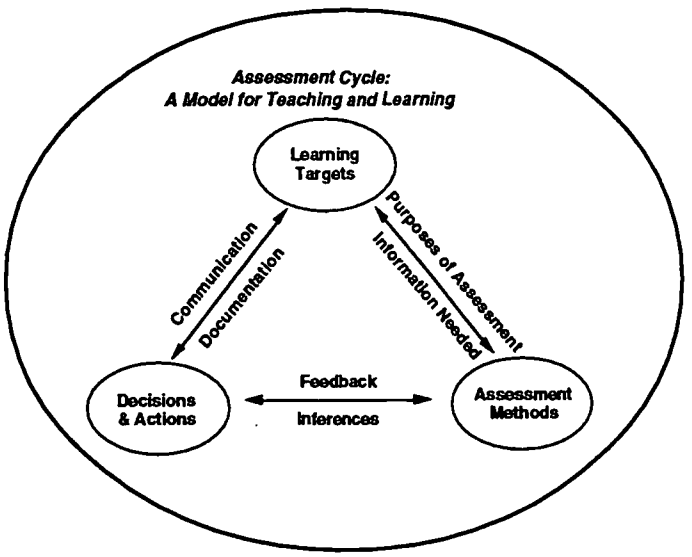
combine, summarize, design, imagine, generalize

Evaluating: Assessing the reasonableness and quality of ideas

judge, evaluate, rate, verify, assess, define criteria

Chapter 3

Using Multiple Assessment Strategies



Because teachers always want to assess students' understandings and abilities to apply content that they are learning, they continually seek efficient ways to gather the needed information. They recognize that some assessment strategies are easier to use than others. Some assessment methods may be easier to administer and to score. Likewise, some methods/types of assessment are easier to create than others. Other methods of assessment are difficult to carry out but rich in the information they yield. Some assessment strategies are more appropriate for young students or those for whom English is a second language. Some assessments are samples of performance and may require high levels of inference. Other assessments evaluate the learning targets directly. In other words, there is not one "best" way to assess students' learning.

There is not one "best" way to assess students' learning. Teachers need to use a variety of assessment methods so that they have the information they need to make decisions.

In deciding how to assess students, teachers match what they are assessing with an appropriate strategy to give them the

information they need effectively and efficiently. They consider how they are going to use the information – to plan tomorrow's lesson, to assign a grade, to check on an individual student's understanding, and so on. As teachers put together the "what students need to learn," the "how to find out about their progress," and the "how to use information about where students are," assessments begin to influence teaching and learning in positive ways.

Teachers use assessments to
- Find out what students already know,
- Make instructional plans,
- Monitor individual student's progress,
- Assign grades, and
- Summarize students' accomplishments.

If, for example, a teacher is near the end of a unit and trying to determine if students have learned some factual information about the American Revolution, multiple choice or short fill-in-the the-blank questions are efficient ways to check for this knowledge. If, however, the teacher wants to find out about students' understanding of the issues that influenced each side to take up arms, open-ended or discussion questions may be the more appropriate assessment strategy.

Assessment strategies

There are five basic categories of assessment strategies that teachers frequently use. Within the categories, there are several different methods. Each assessment strategy has advantages and disadvantages. Some are more appropriate for different learning targets than others, but each method offers some kind of evidence of student learning. The main categories are

- Selected response (forced choice),
- Student written responses,
- Performances,
- Conversations, and
- Observations.

Selected response assessments include multiple choice, matching, and true-false questions. Written work of students (student constructed responses) includes responses to short-answer and open-ended questions, essays and research papers, reports from investigations and book reviews, and journals and learning logs. Another category of assessments includes original performances, products and projects, and responses to complex, "real world" performance tasks. Conversations are interviews, informal discussions, oral questions posed to the class as a whole or to individuals, Socratic seminars, and student conferences. Observations include both planned and informal observations.

Take Time to Reflect

- What is the purpose of most of my assessments?
- What do I already know about different assessment methods?

Selected Response Assessments

Selected response assessments (forced choice assessments) are efficient ways to assess content knowledge. Multiple choice, true-false, and matching tests are examples of selected response items. Short answer fill-in-the-blank questions also are sometimes included in this category, since they are likely to be scored right or wrong.

Since there are aspects of knowledge (information) involved in all that we do, it seems appropriate to use this form of assessment along with other strategies, especially when the learning target is factual recall. However, selected response questions are not limited to assessing factual information alone. Well-constructed, these types of assessments might ask students to compare or contrast, to make predictions based on data, or to summarize main ideas. Many standardized tests, given for accountability purposes or to compare students with norms, are able to sample a broad range of knowledge by using selected response items.

The advantages are obvious. Multiple choice, true-false, and matching tests are quick and easy to score by hand and relatively inexpensive to score by machine. Because answers are marked right or wrong, they can be scored by someone other than a teacher. Multiple choice items are generally considered to be more objective measures. When subjectivity does exist, it may lie with the writer who decides, for example, in multiple choice formats, what the one correct answer is.

Another advantage is the availability of items. Most textbooks provide multiple choice tests with their books. Item banks and computer management systems give teachers access to an array of multiple choice items that are

Selected response assessments include
- Multiple choice,
- Matching, and
- True-false.

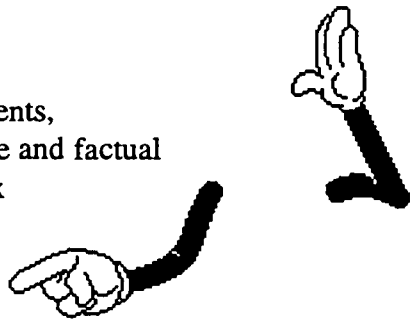
Written responses of students include
- Responses to short-answer and open-ended questions,
- Essays and research papers,
- Reports from investigations and book reviews, and
- Journals and learning logs.

Performances include
- Products and projects,
- Debates, and
- Responses to complex, "real world" performance tasks.

Conversations include
- Interviews,
- Informal discussions,
- Oral questions posed to the class as a whole or to individuals,
- Socratic seminars, and
- Student conferences.

Observations include
- Planned observations,
- Informal observations.

Selected response assessments:



Also known as objective tests or forced-choice assessments, selected response is frequently used to assess knowledge and factual information. Well-constructed, they can assess complex understandings. Students may be required to integrate knowledge and interpret information through selected response items.

In this type of assessment, students must choose from the options presented by the test maker. The thinking involved in *selecting* a correct option may be very different from that involved in *constructing* a reasonable response.

Disadvantages are that this type of assessment is often used as the sole source of information about students' learning. Students, especially creative ones who do not see things like everybody else, may not have ample opportunity to demonstrate all that they know and can do. Teachers see only answer choices and do not have an opportunity to see the thinking (processes) students use to arrive at the answers. Selected-response assessments can encourage guessing; true-false tests in particular allow a student a 50-50 chance of answering successfully with a guess.

Definite advantages of this type of assessment are that they can be quickly machine scored and are marked right or wrong, allowing someone other than the teacher to do the scoring. Teachers and others can use bubble sheets and scan students' responses quickly.

professionally written. These are helpful in surveying students' grasp of large amounts of information or when teachers want selected response items to include with their own questions in assessments they create within the classroom.

The samples on page 47 illustrate different types of selected response and fill-in-the-blanks assessments that have a single possible correct response for the blanks.

There are, however, disadvantages. Teachers are not able to see the processes students employ to arrive at an answer choice nor are they able to know the reasoning students are using. Students who are poor readers may have difficulty making appropriate answer choices not because they do not know the content but because of the wording of the items.

Selected response formats are not usually the best assessment methods for young children. Young students are very influenced by their personal experiences and want to "tell the teacher" much more than the answer choices allow. As a result the child

Sample Selected Response Assessments

Match the decimals with equivalent fractions or percents.

____ .05	A. 1/2
____ .75	B. 5%
____ .25	C. 3/4
____ .5	D. 3%
____ .03	E. 1/4

Choose the best synonym for the *italicized* word in the sentence below.

This change should e*xpedite* the settlement.

A. facilitate

B. expand

C. acerbate

D. eliminate

The _____ is the longest river in Egypt. Throughout history this river was called the "life-line" of Egypt. The river flows from (give directions) _____ to _____.

Fe is the symbol for iron. Write the symbol for these elements:

sodium _____

calcium _____

Circle to indicate if each statement is true or false.

True	False	1. Movement of rock and soil to other locations is one result of erosion.
True	False	2. Gravity plays an important role in erosion.
True	False	3. Because erosion depends upon stream slope and volume, only rapidly moving streams cause erosion.
True	False	4. Long-term effects of erosion include a gradual leveling of the land.

may not respond at all or may choose indifferently because none of the choices seem appropriate.

What is important for teaching is that the assessment method be appropriate for

- The purpose of the assessment,
- The students who are being assessed, and
- The way in which the assessment data will be used.

Sometimes selected response assessments give the impression that young students know more than they really understand. They may have memorized factual information but not be able to apply the knowledge. Likewise, they may have more knowledge than they demonstrate through their answer choices. They tend to be very literal and may focus on aspects of the prompt that lead them away from the question being asked.

They may bring their own experiences to the pictures on the page rather than attend to the questions. For these and other reasons, primary teachers may wish to limit their use of this type of assessment.

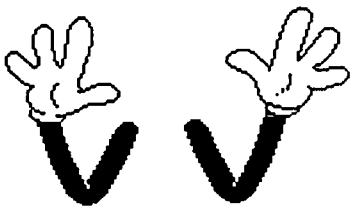
Take Time to Reflect

- What kind of learning targets have I assessed with selected response questions?
- What did I find out about my students?
- What else do I want to know?
- What skills do students need to be successful with selected response assessments?

Students' Written Work

Students' written work (constructed responses) provides a valuable source of information about what they understand. Drawings with captions, solutions to problems, essays, reflective writing, and responses to open-ended questions are examples of written records which help teachers decide what to do next. They also assist teachers in monitoring students' progress toward achieving learning targets and are useful in documenting progress over time.

When assessing through students' written responses, it is clear that age and fluency with language are extremely important. Fluency with language and the ability to express ideas in writing



Students' written work:

Many classroom assessments are based on students' written work (constructed responses). This form of assessment is perhaps the most frequently used way to gather "evidence of learning."

Students' written work includes essays, answers to open-ended questions, labeling, filling in the blanks (classified as selected response by some authors), constructing webs, charts and graphs. Journals and learning logs, spreadsheets, and students' self-assessments can be considered constructed responses since most are the result of a prompt (question) posed by the teacher. The prompts may be very specific or they may be very open-ended.

are both goals and necessities. For example, in primary years students are emerging as writers. Teachers may need to add additional comments on their papers to clarify temporary spelling or explain ideas that are partially recorded. However, given experiences in recording their thinking and having their ideas valued, students do become better at written responses.

While student constructed responses often assess recall, basic comprehension, and simple application, they can also give teachers opportunities to assess critical thinking and reasoning and make judgments about the depth and complexity of students' understandings. Essays, reflective writing, and responses to open-ended questions reveal the diversity of students' thinking. These questions may provide opportunities for students to respond to prompts that have multiple, appropriate solutions which need to be "backed up" in the discussion or to respond with the "right answer" but describe one of several valid solution paths.

Some questions may require only brief answers or short paragraphs; other prompts require longer discussions. Verbs like list, describe, identify, compare, explain, predict, and evaluate are often used in questions for students' written responses. Frameworks for thinking and reasoning are helpful resources for teachers interested in asking higher order questions. Because teachers are often able to assess

Tips for Improving Written Responses

- Model responding to questions
- Provide adequate time to respond
- Encourage the use of diagrams, charts pictures, graphs to support the commentary
- Respond with feedback that gives students specific ways to improve writing
- At times, allow editing and revising
- Scribe for students unable to write for themselves

When students write about their thinking...

When students write about their thinking and their opinions in journals and learning logs, both teachers and students benefit. Students learn to reflect on what they know, what they do not understand, and what questions they still have. In this way they are self-assessing.

If reflections are to be honest, the many journal entries should not be graded as right or wrong. When students use journals to express their feelings, they have a private way to communicate with the teacher. What is appropriate to evaluate may be the prompts that ask students to state an opinion and to give reasons to back up their ideas. The grade comes from how well their position is supported, not the stance they have taken.

Reflective writing is not easy for many students. Teachers need to emphasize the importance of being able to communicate in writing. In primary classrooms making a class journal on large chart paper is a way to model with the group what teachers expect students to do independently at a later date.

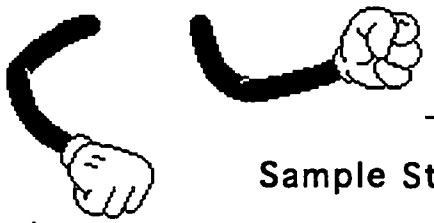
Students need to be clear about what their journal assignments are. That is, unless there is a definite purpose for writing, students may feel that it is a waste of time. Likewise, if the teacher is not the audience for a particular reflection, students need an indication of who they are writing to. Teachers need to show that they value the students' ideas by responding to students in the journals. One strategy is to read five or six journals a day rather than trying to respond to a whole class set at once.

Students' journals may be small booklets that are two or three sheets of paper folded and stapled. In upper grades small spiral notebooks provide space for students to include drawings and diagrams along with their narratives.

If simple knowledge is the learning target, selected response may be the most efficient assessment strategy. When assessing integrated knowledge, student constructed responses may give a truer picture of students' understandings.

conceptual understanding to a greater degree with written responses than with selected responses, classroom assessment should include opportunities for original works.

There is more subjectivity in scoring these assessments than in scoring selected response items. Having specific criteria for assigning scores is an important strategy for consistency in scoring. Teachers and students need to be clear on the content targets and what high quality performance looks like. When students receive scores or grades that they do not understand or are lower than they expect, relating their work to the criteria can provide feedback for greater success on future work as well as an explanation for the current scores. Thus, rubrics help teachers, students, and families understand the expectations and learning progress.



Sample Student Constructed Response Questions

- After reading T*en Black Dots by* Donald Crews, the teacher posed this question for her students. "If you were going to make your own black dots book, how many dots would you need? Be sure to let me know how you figured this out."

- In another classroom, the teacher posed a similar question, "Including the hunter, how many animals were there?" after the students had read *1 Hunter* by Pat Hutchins.

- *A Spanish II teacher gave her students the following prompt:*

  Your family will be hosting an exchange student from Peru. Since the student does not speak English, write a letter in Spanish introducing yourself and your family.

  Tell the student about your daily routine, your hobbies, interests, and your school.

- Christianity, Islam, Hinduism, Buddhism, and Judaism are called the five major religions of the world. Describe each religion, giving examples of how they are similar and how they differ.

- *In what ways is the job of a television news announcer like that of a town crier in the Middle Ages? In what ways is it different?*

- People who have a common race, language, religion, and cultural background may be considered to be an ethnic group. Briefly describe these groups: Aborigines, Basques, Kurds, and Lapps.

- *List the seven continents in order of size from largest to smallest.*

- Jamal claimed that he ate more pizza than his friend Cher because he ate one-half of his pizza and Cher ate one-third of the pizza she had. Cher said she thought that Jamal was wrong because she ate the most pizza. Which student do you agree with and why?

- *Gulliver had many adventures in his travels. Pretend you were a Lilliputian who travels to North Carolina and write about one adventure that could take place in your home town.*

Take Time to Reflect

- Am I encouraging students to explain their thinking and develop writing skills that they can apply in all content areas?
- When assessments include essay or discussion questions, are the students clear about my expectations for full and complete answers?

The advantages of using students' written work to assess their learning are several. Teachers are able to assess a variety of learning targets and complex performances. Because the work is created by the students, there are opportunities to gain insight into students' thinking and reasoning. Written assessments such as essays and responses to open-ended prompts take less time than complex performance assessments and yet may yield much of the same information. There is also the advantage of having the original document to save in portfolios or working folders.

Disadvantages of assessing learning through students' written work relate to the time it takes for scoring, the difficulty in scoring different responses students' may give, and issues related to clear, consist application of expectations (standards). *How much consideration (flexibility in scoring) should be given to different students versus holding every student to the same criteria for proficient work? What should the flexibility be for students who have special needs related to language?* Students must understand what is being asked of them and have sufficient fluency in the language to respond if student constructed responses are to give accurate information about achievement.

These are important questions that might be discussed in department or faculty meetings. For further discussion, see Stiggins (1997), chapter 15.

Performance Assessments

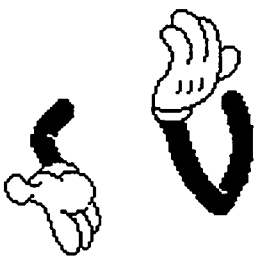
Performance tasks, presentations, products, investigations, projects, and original creations are important ways in which students demonstrate their abilities to make connections and to apply their skills and understandings. These assessments may take several days or even weeks to complete. They are often referred to as "authentic assessments" because they mirror expectations that students will encounter as adults. Sometimes

classified as complex applications and other times separated into different assessment categories, these assessments share the similar aspect of students making connections and integrating their learning.

Performance assessments mirror real-life experiences and may involve students working together in problem solving situations that require working through multiple steps in a process, doing research, and making many decisions before completing the task. Since working together with other individuals is a goal that is valued in the workplace, performance assessments offer opportunities for students to be evaluated with a rubric that judges affective goals as well as cognitive ones. Respecting individual differences, learning to identify and delineate responsibilities, listening to others, sharing ideas, and learning to compromise are important skills that may be evaluated as students complete performance assessments.

Good questions are fundamental to all assessments.

Performance assessments:



Performance tasks and presentations are often real world challenges and problems that require students to "pull things together." Students must integrate concepts, skills, facts, reasoning and problem solving to complete these assessments. Debates, investigations, and presentations that include oral presentations and written documents are examples.

These assessments may require the teacher and student to be present for the evaluation process. The tasks and performances are most often evaluated on multi-dimensions using rubrics. Time to implement and to score may be disadvantages, but being able to assess complex learning is a definite advantage.

Proponents of these methods of assessment point out that the assignments mirror and measure what we say we value in education. They involve higher-order thinking and require that students be active workers not passive test takers. They are said to be accessible to students with different learning styles, different backgrounds and experiences, and varying abilities. These assessments look more like learning activities than traditional tests. Performance assessments range from tasks that can be completed in short periods of time to involved projects which take a number of days to complete.

Criteria for Performance Tasks for Instruction or Assessment

- Measures important content - the task should represent a "big idea" in the curriculum and should mirror real life
- Engages learners - the task should be interesting to students and should provoke thought, enthusiasm, and persistence
- Enhances learning - the task should promote continued student learning, allowing the student to be a decision-maker and an active worker
- Is equitable and feasible - the task can be accomplished at school by all students and is developmentally appropriate

Often students are able to become part of the process of clarifying the performance criteria and defining scoring rubrics. These attributes, which are viewed as advantages by some teachers, may be disadvantages in the eyes of others. Performance assessments do require a large amount of time to create and administer. Likewise, scoring of performance tasks can become very subjective if teachers are not explicit about the standards and careful in determining appropriate criteria for different levels of achievement.

Because these assessments have multiple dimensions, analytic rubrics (see pp. 82-86) are likely to be helpful in communicating expectations for quality work and successful achievement of the assignment as well as scoring the final results.

Since class time and teacher preparation time are limited, teachers must weigh their expertise in designing these assessments and the benefits of using complex performance tasks with the information (evidence of learning) that they produce. There is no question that quality tasks, well-implemented and well-scored, add an important dimension to the evaluation of students' learning. For this reason, teachers may decide to begin with a published task that has been used by others and has scoring criteria in place.

Take Time to Reflect

- Do my assessments ask students to demonstrate their understanding of concepts?
- Do I ask them to apply their knowledge to accomplish complex tasks accurately and appropriately?
- Do students have a clear understanding of how their performances will be evaluated?
- Am I building a file of quality performance assessment tasks?

Sample task: As the culminating activity in this public speaking class, you are to research a community problem and prepare a 5-7 minute speech discussing the problem and making recommendations for how it might be resolved. You will be evaluated on the following criteria:

- accuracy of the factual information,
- reasonableness of your solution,
- clarity of your main arguments, and
- persuasiveness of your presentation.

Quality tasks motivate and engage students as well as require them to apply conceptual understanding and skills.

Sample task: You have been asked to make recommendations to the Cafeteria Manager about the kind of extra items her staff should be selling each day at lunch. She is interested in the nutritional value of the items you suggest, their appeal to students, and the profit the cafeteria might make on these items. She is expecting a written report, complete with charts and graphs and a description of how you conducted your research. Your final report and recommendations are due in one week.

Sample task: Your team is to design a web page about your school that can be accessed for new residents by the realty companies who are trying to sell houses in your school district. The site should include information about the school's academic program, student achievement, and extra-curricular opportunities. You will be presenting your product to the review panel in three weeks. The panel is comprised of four students, the assistant principal, the PTA president, and a representative of a local real estate company.

Page 56 illustrates a task that integrates mathematics and social studies.

Preparing for Performance Tasks

A tale of two high school teachers at work

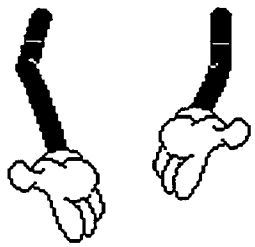
Two high school teachers, one in mathematics and one in social studies, decided to work cooperatively to prepare students to respond to more complex assessments. Both teachers were interested in using a performance task as part of their semester grades, but students had had little experience with this type of assessment. The teachers felt that having students learn how to analyze performance tasks, help create criteria for high quality performances, and self-assess was important for their students.

The task, printed on the next page, was introduced in the ELP classes (Economics, Legal, and Political Systems). Students discussed the information they would need to complete the assignment and where they could get the data. Students brainstormed resources, and time in the media center was scheduled for the classes.

The students and the ELP teacher talked about how grades would link to the agreed upon criteria for clear and persuasive presentations. The accuracy of their data would also be evaluated. The teachers supplied maps of North Carolina which showed the 100 counties.

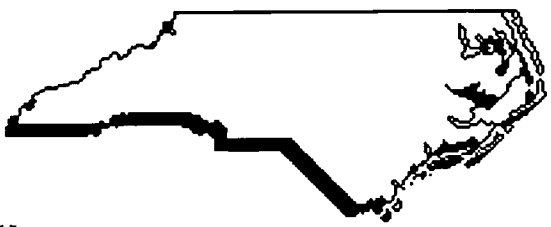
In their mathematics class the students talked about what mathematics would help them complete the task. The teacher asked them what they already knew and what they needed to learn. He challenged them to think about the kind of algorithm they might develop to solve the problem. The teacher explained that he would be grading students on the appropriateness and accuracy of the mathematics they used and the charts and or graphs they created. The class worked with the teacher to create criteria for grading these aspects of the performance task. They included justification in the rubrics.

Since there are 100 counties and 50 Senate districts, several students suggested that the Senate districts would be easy to assign – each district will be two adjacent counties. The teacher reminded the class that unlike in the U.S. Senate, in North Carolina every district, whether Senate or House, must represent approximately the same number of people. He suggested that they choose two pairs of adjacent counties to see if the one Senator per two counties suggestion would work. To further clarify the equal population guideline, the teacher asked students to experiment by creating two House districts, one in the far west and one in the central part of the state, and to compare the populations of the districts they drew.



This is the prompt that was given to the students.

Reapportioning the North Carolina Legislature



You have been appointed to a task force to reapportion the North Carolina Legislature. Your assignment is to draw up and record fifty Senate and one hundred twenty House districts using the 1990 census data. Certain guidelines apply:

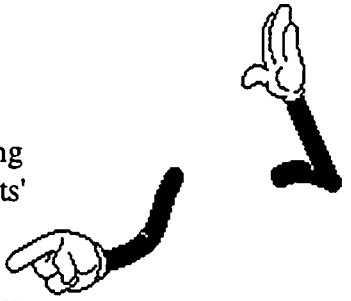
1. Contiguous counties should be placed together to form a district.
2. Or, one county may be divided into several districts but a portion of a county cannot be placed with an adjacent county to form a district.
3. All Senate districts should have approximately the same population.
4. All House districts also should have approximately the same population.

You are to appear before House and Senate leaders in one week to present your recommendations. All of your work must be explained and justified mathematically. The committee expects that your presentation will include appropriate visuals. Accuracy is of utmost importance since your plan may be presented in hearings across the state.

Conversations, conferences and interviews

Often what we expect a student knows and what the student really thinks are different things. Informal conversations, conferences, and interviews provide teachers opportunities to probe students' responses. This information influences the feedback teachers give students as well as the teacher's future instructional plans.

Conversations, Conferences, and Interviews:



Conversations are one of the best ways to assess students' thinking and reasoning because they allow the teacher to probe the students' responses. They are helpful in assessing students' understanding of processes and in finding out the extent of students' factual knowledge and conceptual understandings. Interviews are usually thought of as one-to-one conferencing, while conversations may be with small groups, individuals, or the class as a whole. "Think alouds" are a traditional language arts assessment strategy that exemplifies this labor intensive but rich means of assessment.

Conferences are an important means of providing feedback to students about the quality of their work and their strengths and misconceptions. If both student and teacher prepare for the conference, there are opportunities for each to share thoughts about the quality of the student's work as well as the student's progress toward learning targets.

For young students, or those with special needs, conversations are likely to provide the most accurate information about what the students know. Like observations, however, the assessment strategies of conferencing and intervening have the disadvantage of needing some kind of documentation.

Interviews and conferences with individuals and small groups is time-consuming but needed. Informal discussions take place when teachers stop beside students' desks for two or three minutes to ask or answer questions. A three minute, spur-of-the-moment conversation about a particular issue, task, or problem may provide more understanding of what a student is thinking than much longer written assessments because teachers can tailor their "next questions" based on the specific responses given by the student. These informal conversations are important for the students also, because they can receive specific, immediate feedback as the teacher is gaining insight.

The story below about an informal conversation illustrates how students can perform routines, but may not understand underlying concepts. Often students repeat what we model for them and may give, as Kathy Richardson (1997) says, "illusions of learning." They are mirrors that reflect back what we say and do, but they do not necessarily have the understandings we might assume from their responses. Think about what Timothy does and does not understand.

Let's talk about Timothy...

Timothy's worksheets on finding equivalent fractions had all correct answers. The class had spent several days on fractions, completing exercises in their fourth grade textbook. Now they were finishing the unit by "reducing" fractions. Timothy made an "A" on the weekly quiz . His puzzled looks during class discussions caused his teacher to wonder about his understanding. She related this story:

I paused by Timothy's desk and asked him to explain what an equivalent fraction was. '"It's like this," he replied. "You just multiple the top number and the bottom number by the same thing." Timothy wrote on his paper 1/3 x 3/3 = 3/9.

"Could you draw me a picture or explain in another way?" I asked him.

"No," Timothy said, "but I can give you another example." He wrote 1/4 x 2/2 = 2/8.

"Tell me about reducing fractions."

Timothy responded as he wrote 6/8 ÷ 2/2 = 3/4, "You just divide the top number and the bottom number by the same number."

I asked him, " If I draw a two rectangles that are the same size and I shade one of them to show 3/4 and the other one to show 6/8, will the same amount be shaded?"

Timothy looked at me and then pointed to 6/8. "This one would be more because 6/8 is bigger."

Young students frequently know more than they can record in traditional, symbolic formats. They arrive at answers, but find it difficult to write out the steps they used to arrive at their solutions.

Observations and conversations are likely to give the most meaningful information about these students.

Elementary language arts teachers are familiar with conferences which focus on reading and writing; however, there is a trend for more teachers in all disciplines to use conferences to help students set individual goals, evaluate their work, and review their work habits. Formal conference schedules are a way to allocate time for in-depth conversations. However, rather than schedule conferences with all students, many teachers interview students that are either having difficulties learning what is being taught or that seem to need to move ahead of the class.

Tips for Conferencing with Students

- Begin with a few students when you are convinced that the benefits of one-on-one conversations outweigh the difficulties.
- Keep conferences brief.
- Plan so the rest of the class can work independently.
- Have a specific purpose (topic) in mind for the conference.
- Be prepared with questions to gain insight into the students' understanding related to the topic.
- Have the students bring samples of work related to the topic.
- Use conference time for positive interaction for you and the students.

Required in both formal and informal interviews are effective questions. Many of the same questions that teachers pose to the class are appropriate for individual interviews. These are questions that ask students to explain, to tell why or how, to give additional information, or describe an alternative. Effective questions ask students for more than just the right answer. They encourage discussion and avoid formats that can be answered by a yes or no.

Socratic seminars and small group discussions combine assessment through conversation and observation. While teachers may choose to be primarily an observer, they also are able to prompt and probe students with follow-up questions.

In upper grades teachers may evaluate students through class discussions or oral examinations. This requires very careful listening, and teachers must be cautious not to draw conclusions from brief responses. Frequently, additional questions and probing are needed to fairly evaluate the student's understanding. Likewise, there must be consistent criteria which guide the way students' answers are evaluated and efficient ways to record students' scores. One potential problem with oral examinations is that time may permit only a sampling of what individual students know and our own biases may keep us from providing sufficient opportunities for students to "make their cases."

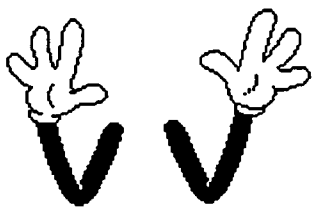
Take Time to Reflect

- Do you have other examples, similar to the story of Timothy, of when students first appear to understand important ideas and you later through conversations discover that they have major misconceptions?
- How do you use conversations, conferences, and interviews to assess?

Observations

Teachers have always engaged in "kid watching." It is one of the oldest means of assessment, and it has been used to monitor students' progress since earliest times. When standardized tests

became popular, observation was deemphasized as a means of formal assessment and only recently has it again become valued as an assessment tool. Observations give teachers information about individuals and about the class as a whole. Teachers are able to observe the processes that students use and are aware when students' verbal and nonverbal behaviors do not match.



Observations and Checklists:

Teachers constantly make judgments and take action based on their observations in the classroom. Observing provides ongoing, unsolicited information about the class as a whole and about individual students. Teachers can see students using skills and applying their understandings to solve problems and accomplish tasks.

An advantage to this method of assessment is that students may demonstrate understanding that is difficult for them to explain in written assessments that are heavily influenced by reading or language abilities. A disadvantage is that teachers must be skillful in making systematic use of observations and in creating checklists and other anecdotal records that become the written documentation of what they see.

Informal observations occur during instruction, when students are working individually or in groups, in the halls, or on the playground. Because different teachers can observe the same behavior and make different interpretations, it is important that inferences that arise from informal observations be based on teachers' strong content knowledge and clarity about what accomplishment of learning targets will look like. Multiple opportunities to observe and interact with students provide important insight into students' thinking and understandings, not just what students can do as memorized processes.

Both verbal and nonverbal clues students give related to their thinking provide teachers with important information.

More formal observations are planned as teachers decide ahead of time what they are going to look for and how they will record the assessment data. These observations are likely to be more useful if teachers develop a systematic way to pay attention to all of the students. This may mean focusing on only a few students each day or looking for evidence of a

Making Systematic Observations

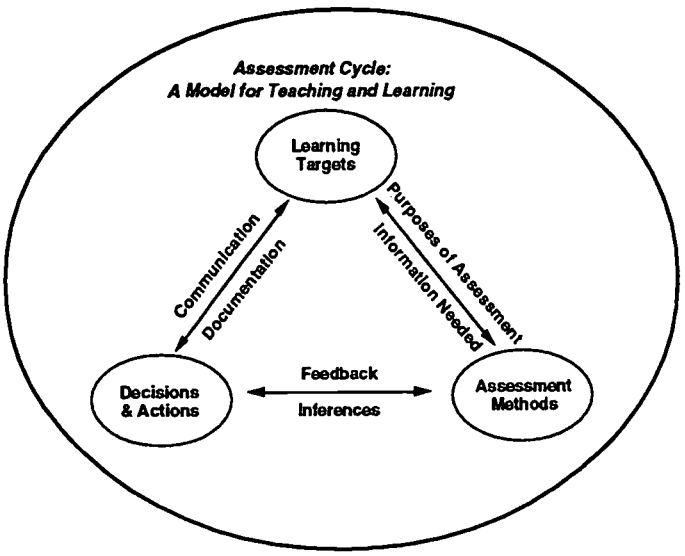
- Plan to observe on a regular basis
- Include all students, though not necessarily all on the same day
- Make some kind of notation of important information
- Observe each student more than once, looking for patterns and consistency

few, specific understandings in all of the students. Noting what is typical as well as what is atypical in students' responses and actions is helpful in making instructional decisions and monitoring the progress of students.

Checklists, note cards, matrices, and notes on preprinted labels are examples of record keeping strategies that arise from observations. These take a number of different formats. Some are developmental; that is, they may be used repeatedly throughout the year to monitor students' growing expertise or they may list related, increasingly sophisticated traits. Other checklists and observational notes may record skills or processes that the teacher is monitoring. Some teachers use one checklist or recording sheet for each student with many factors addressed on the page. Other teachers list groups of students or the entire class on one page and use it to record information about one or more attributes.

Chapter 4

Making Decisions and Taking Action



Walk into many classrooms today and you will find teachers who know a great deal about their students. They have accumulated a lot of information using a variety of assessment methods. Based on how they interpret this information (that is, the inferences that they make), teachers will take any number of actions. They may decide, for example, to reteach, to move on, to regroup for further instruction, to reassess certain learning targets, or to assign a grade.

An inference is a conclusion about students' understanding that cannot be directly observed.

However, before we can talk about decisions and actions, we first need to talk about how the information we derive from assessments is interpreted. When a teacher scores a student's work, it is not the grade that gives the teacher rich information about what the student knows. Rather it is the feedback the teacher gets from studying the actual work of the student. Teachers ask themselves

- *What has the student learned?*
- *How well does the student understand the material?*

- *Is there evidence of confusion or misunderstanding?*
- *How well is the student applying and using the information?*
- *Has my assessment given me a sufficient understanding of the student's thinking and reasoning?*

Thinking about what students know and how they learn gives meaning to information that is gleaned through assessments. We are able to make good inferences when we put all of this information together. This includes our knowledge about how students learn, what the learning targets are, and the specific information from the students' performances on the assessments.

Quality Inferences

Two questions must be considered for quality inferences. These are critical to the process of using inferences to make decisions:

- *Do I have enough information for the inferences that I am making?*
- *How directly am I able to assess the learning targets?*

As every teacher knows, a student may perform differently on different days and in different situations. Therefore, before making the decision that a student has adequately mastered the learning target, teachers need multiple evidences over an extended period of time. For example, students may perform very well just after the close of a unit of instruction, but in two weeks (or two months) their level of performance is dramatically lower. Likewise, students may be able to solve similar problems but are not able to use the same information when it is presented in a different context.

One way to decide if there is sufficient evidence is to consider the consequences of the inferences. Judgments that influence the lesson for tomorrow need less evidence than judgments that go home on report cards or that influence whether or not a student goes into a special program or on to the next grade. Breadth and depth of evidence should always be determined by the importance of the decisions that are being made and the opportunities students will have, or will not have, to master the learning targets.

Assigning meaning to students' words, actions, and products is perhaps the most difficult part of assessment. However, teachers must deal with students' misconceptions as well as their strengths if students are going to be successful. If decisions are made from too little evidence or misleading evidence, teachers may not plan the necessary classroom experiences for the students to change their thinking.

The second consideration, that of directly assessing a learning target versus having to draw more conclusions from less direct evidence, is more difficult to get a handle on. For example, if the learning target is to be able to speak standard English, teachers can assess it both directly and indirectly. In conversations, a teacher hears the student's use of grammar and choice of sentence structure. Another less direct way to assess this learning target is to ask the student to edit text written by someone else. A student's ability to identify incorrect grammar in another student's paper may not transfer completely to the ability to use correct grammar fluently.

In the first example, the ability to speak standard English is observed directly and requires little inference (that is, it is a low inference measure). But in the second example, the teacher must infer the student's ability to speak standard English (a higher inference measure).

High Versus Low Inferences

Some assessments–i.e., students demonstrating how to do something–allow more direct observation than others. As illustrated above, other assessments are less direct and require teachers to infer rather than find out directly. Different ways of assessing are measures which range on a continuum from high inference to low inference.

High inference procedures measure less directly the understandings that the student may or may not have. For example, standardized tests are higher inference than direct teacher observations and conversations. From the test scores, based on number of items correct, inferences are made about students' knowledge of content being assessed. Among standardized tests, assessments that measure general constructs, like intelligence, are higher inference than achievement tests in specific content areas.

Student work products are likely to vary in level of inference depending upon the task and the learning target that is being measured. While high inference procedures are related to the learning targets, they are less likely to get at the *specific* concepts and/or understandings held by the student. These less direct measures are often used in summative decisions and/or higher stakes decisions where there is a desire for the assessments to be consistent across classrooms and schools.

When an assessment has students demonstrate an application or give factual information, it is direct evidence that needs little interpretation. For example, students can suggest synonyms or label bodies of water on a map. They can mix varying proportions of paint to make different hues, or they can type a business letter correctly. Having students apply a procedure and explain the steps, telling why they did what they did, gives teachers opportunities to assess directly skills and processes. The level of inference is low because the evidence is *direct.*

Inaccurate inferences and assumptions are likely to lead to poor decision making and false impressions of students' achievements.

Assessing students' thinking, reasoning, conceptual understandings, and problem solving is more difficult than assessing factual information and procedures. Measures are likely to be less direct and require higher inferences unless teachers are able to find out about students' thinking processes as well as their products. Teachers ask questions such as *How did you reach that conclusion?* or *Why did you choose this way of presenting your information?* to try to measure a directly as possible complex understandings. Pictures drawn by young students, written explanations, and performances are frequent choices for assessing these types of targets. Direct observations and individual conversations are powerful strategies. Teachers need to plan a systematic way to incorporate these strategies into the classroom routine rather than viewing these assessments as add-ons.

The challenge for teachers is to make sense of what they learn about individuals and the class from their on-going assessments. That is, because direct assessment information and the inferences teachers make from indirect evidence are the basis for many of their decisions, teachers need as accurate a reflection as possible of what students know, understand, and can do. Inaccurate inferences and assumptions are likely to lead to poor decision-making and false impressions of students' achievements.

Classroom assessment employs many direct (low inference) measures that are most useful for day-to day instructional decisions as well as cumulative data about student knowledge and skill. Different levels of inference may be found within one type of assessment depending upon the learning targets and the way in which the questions are framed.

Interpreting Assessment Data

Take Time to Reflect

- Do I understand the difference in direct and indirect assessment?
- Am I using these appropriately?

Valid inferences require informed judgment on the part of teachers. Knowledge of the content, what is relevant to the learning target, the quality of the assessment data, and variables related to individual students are all components that can influence the validity of inferences. Low inference measures, such as students' demonstrations, conversations, and observations, are most likely to yield information about students' understanding of the content. Therefore, on-going assessment appropriately employed by teachers has the potential to enrich students' overall achievement. Effective feedback for students and focused instruction can grow out of valid inferences.

Making judgments about students is best done with a convergence of evidence rather than a single assessment.

It is not a simple task to attribute meaning to students' words and actions to serve as indicators of student learning. It is crucial that the interpretations are as accurate as possible. Important decisions about learning and teaching arise from the interpretations. Therefore, teachers must think carefully about the content they are teaching, the continuum of understandings and skills related to the learning targets, and the information they have gathered from the assessment.

Teachers ask themselves:

- *Do the data indicate that students (individuals as well as the class as a whole) are ready to move on?*

- *Is there a misconception that needs to be addressed?*
- *Do students appear to understand the concept but need more practice in applying their knowledge?*
- *Which students need more time to explore the ideas?*

Notice in the example that the student used the long division algorithm to correctly compute. The story that she wrote, however, does not match the computation. What does this tell about the student's understanding of division?

When asked to write a story problem (an activity that the class did frequently) to illustrate \$8.24 ÷ 4 and then to solve the problem, one fifth grader wrote the following:

Mary and her four friends found \$8.24. They spent \$12.64 at the store. How much did they have left?

$$
\begin{array}{r}
2.06 \\
4\overline{)\$8.24} \\
-8 \\
\hline
2 \\
-0 \\
\hline
24 \\
-24 \\
\hline
0
\end{array}
$$

Since many decisions about students are made in the classroom, it is helpful for teachers to be aware of issues related to bias. Research and experience in the classroom have verified the "self-fulfilling prophecy" that students tend to perform as teachers expect them to. Likewise, teachers tend to see and hear what they are expecting. Thus, making judgments about students is best done with a convergence of evidence rather than a single assessment.

Because students learn different content at different rates and in different ways, on-going assessment will suggest adjustments to guide teaching and flexible instructional groupings. As the school year progresses, teachers gather more information and continually reevaluate the inferences they have made about individual students and the class as a whole. This process–a continuous cycle of setting learning targets, teaching, assessing students, making inferences and adjusting plans–is what classroom assessment is all about.

Quality Evidence

One difficult aspect of any assessment is determining what constitutes quality evidence. That is, is the evidence at hand

worth counting? Because decisions are made from the inferences that result from interpreting assessment data, it is important to think about the issues that surround quality evidence. Teachers must be able to continuously evaluate their own procedures in order to improve the validity of the interpretations made by them about students' achievements.

Consider these issues:

• *Are you able to clearly describe the learning target? Is the target simple or complex?* Learning targets must be congruent with what students in a given class/grade are expected to achieve. When the target is complex, such as developing an understanding of place value, a wide range of instructional activities and opportunities are needed for students to develop the concept and its applications. Assessment of complex targets such as this example must consist of many sources of evidence in a variety of contexts and formats over time.

• *Can you articulate clear and appropriate criteria for achievement of the learning target?* A deep understanding of the specific content and the important outcomes that relate to a learning target are essential. If teachers do not have a clear vision of what it looks like when students have or have not achieved the goal, they may unintentionally be teaching students to hit the wrong target. They may assume that students have learned the target when they have not; or they may assume that students have not learned the target when they actually have.

• *What is the scope and purpose of the assessment?* Scope relates to the purpose of the assessment in that deciding what to do tomorrow or next week requires less evidence than summarizing students' achievement at report card time or the end of the year. The higher the stakes (consequences), the more varied the evidence and the greater the amount of time expended in gathering it need to be.

• *Do students have opportunities to demonstrate their thinking and reasoning and to show all that they know and can do?* If assessment questions are not clear to students, if language is a barrier, or if the environment causes anxiety, students may not be able to communicate what they have learned. Classrooms in which instruction and assessment are closely

Criteria for Quality Classroom Assessment

- Important and clear learning targets
- Appropriate methods of assessment
- Reasonable, reliable implementation methods
- Public criteria for achievement
- Fair opportunity for demonstrating learning
- Valid inferences (measuring what you intend to measure)
- Useful, constructive consequences

linked and where assessment becomes a part of classroom activities without obvious breaks may help address this issue.

Take Time to Reflect

- Some mistakes that students make come from a lack of information. At other times they reflect a lack of understanding. When do I give information and when do I allow or help students correct their own misunderstandings?
- In my classroom am I meeting the criteria for quality assessment?

Different Kinds of Decisions

Not only do teachers make thousands of decisions each day, they also make numerous types of decisions. Let's look at the different kinds of decisions teachers make:

- *Decisions about what students already know,*
- *Decisions about what to teach next and how to teach it,*
- *Decisions while monitoring students' learning,*
- *Decisions which encourage student self-assessment,*
- *Decisions at the end of instruction (summative decisions such as grading or promotion), and*
- *Decisions about programs.*

Decisions about What Students Already Know

For classroom diagnostic assessments to provide the most useful information, they are likely to be focused on just what a teacher is planning to teach in the next weeks or in the next unit of instruction.

Data from assessments that are administered for diagnostic purposes identify starting or restarting points for instruction. These assessments, sometimes formal written assessments and other times informal questions and activities, yield specific information about individual students' understandings related to what teachers are planning to teach or reteach. They help teachers plan instruction in ways which build on what students know and can do and often identify misconceptions that teachers need to address.

This is in contrast to assuming that the class as a whole needs

to begin at square one or reteaching what many students already understand. Teachers use the information from their observations and assessments to identify appropriate instruction. Since students bring different experiences to the class and therefore have different understandings, diagnosis addresses the different strengths and needs among individuals in the class.

Diagnostic assessments identify students who already know a portion of what is to be taught and need challenges to broaden their understandings. Teachers also identify those students who have particular needs that may hinder them from being successful without additional support. Based on pretest information, teachers sometimes form temporary groups to provide these different learning opportunities.

Decisions about Instruction

Each day teachers make hundreds of decisions. These decisions are both during and after class. Moment-by-moment decisions are made many times during a lesson when teachers rephrase a statement or question. Sometimes they stop and reteach or realize that students do not need the planned practice because they are ready to move on. These are examples of instructional decisions that are made from small bits of feedback that teachers get through interacting with students.

Assessment focuses on the collection of data. Using those data to make decisions in the classroom on a day-to-day basis helps to create a learning environment where the work of the teacher and the student is highly productive.

Later teachers read students' written work or review what happened in class that day or reflect on what they have discovered about individual students' thinking and reasoning. From this assessment information they decide whether to continue with the lesson plan they have made for the next day or to change what they are going to do.

The nature of classroom assessment means that not all decisions are large ones. There are times when teachers "file away" information they have gathered because it confirms plans they have already made or because it is inconclusive. Deciding to gather more evidence before making decisions about moving on is a frequent result of classroom assessment. Teachers continue teaching and look for information to clarify what the students have learned. Meanwhile, they provide opportunities for students to investigate, practice, explore and

discuss. They ask the class more questions and look for information that will help them decide the next tasks. This gathering of evidence is ongoing classroom assessment.

Knowing intermediate steps along the way toward accomplishment of learning targets helps teachers make decisions about their instruction. Reteach, provide more time for practice, use more or less difficult materials, and group students for a few days of differentiated lessons are among the choices teachers make. If misinterpretations about what a student knows occur, there will be opportunities to gather additional information and make different decisions about the appropriate instruction for the student. For this reason classroom assessment is not usually considered "high stakes." However, if instructional decisions are not based on a clear picture of the learning targets and what students understand and can do (e.g., an understanding of what the content is, what achievement may look like, and how the learning targets fit into the larger picture of what is important for this grade and this discipline), students are not likely to have appropriate opportunities to learn.

It is not a dramatic leap into using a greater variety of assessment techniques that is likely to improve the quality of education. Rather, it is the thoughtful shifts that teachers make in their thinking about what, how, and why they are assessing.

Another caution: it is not a dramatic leap into using a greater variety of assessment techniques that is likely to improve the quality of education for students. Rather, it will likely be the small, thoughtful shifts that teachers make in their thinking about what they are teaching and what evidence they need to be certain that their students have the level of expertise and fluency that is desired. As teachers make a conscious effort to choose assessment strategies that give them more information about what students understand and can do, their decisions are likely to lead to more focused instruction (what to teach) and learning opportunities (how to teach) for students.

Decisions while Monitoring Students' Learning

Part of the responsibility of being a teacher is making every effort to determine where individual students are in relation to the goals and learning targets for the course. This responsibility goes beyond assigning grades to work that students do. It requires that teachers design tasks, formulate questions, and talk with students to understand what they know as students work toward achievement of the targets. It means that teachers must sample students' performances adequately for the

decisions they make. The responsibility also means that teachers are able to identify the performance level of students' work and give feedback to students that will enable them to improve.

This is not to imply that teachers must use separate assessments for the purpose of monitoring students' progress. With the same information that teachers gather through the ongoing assessment that helps them plan for instruction for the class as a whole, they have information to monitor individuals and provide feedback to students. These decisions about what and when to give feedback are very important.

A tenth grade biology teacher assigns "completion exercises" several times a quarter. These exercises are designed to help her to assess students' understanding and make instructional decisions. She does not want to assign grades because she feels students are "in progress" toward the learning targets and are not yet ready to be measured against the standards. To encourage students to complete the work, at the end of a grading period she gives one grade based on the number of completion exercises the students have turned in.

Decisions which Encourage Self-Assessment

Many times observers look into classrooms and see teachers working very hard. Often, observers report, students are hardly working. At every level teachers are concerned with unmotivated and irresponsible students. This appears to be especially problematic in upper grades. Unless an assignment is graded, teachers report, many students will not even attempt to complete it.

While classroom assessment is not likely to solve all of these problems, it may be one way to address student apathy and encourage students to become more involved in learning opportunities. Motivation has always been an issue that schools have tried to address in numerous ways.

In classrooms that have a strong emphasis on self-assessment, there are different roles for teachers and for students. Teachers move back and forth between directing (instructing) and guiding. They facilitate as frequently as they direct. This is not to say that "anything goes" and teachers are not "in charge." Providing students with more opportunities for self-direction means that teachers have established a classroom environment where students know how to be more responsible. It means that the teacher's role for establishing clear content and performance standards and in monitoring students' progress is more important than ever.

Students are likely to become good at self-assessment only when learning targets are clear. If classroom assessments are

varied so that they have opportunities to demonstrate their understandings in nontraditional as well as traditional ways students are more likely to be motivated to take greater responsibility for their learning. They need to be able to look at models of proficient performance, to compare their responses with rubrics that describe criteria for high performance, and to see for themselves what they must do to improve.

There are many well-established examples of students' learning from each other in schools today. Students critique each other's drafts in writers' workshops; they give feedback that is informative and tells other students what they need to do to improve. As students present solutions to problems in the math class, they clarify their own ideas and model alternative ways of thinking and reasoning as they explain their ideas to others.

Before grading a set of papers a teacher might quickly read through the entire and select samples that clearly indicate different levels of performance. These become the anchor papers that will allow grading to be easier and more accurate.

Creating a classroom culture that motivates students to take greater responsibility is a challenge. It requires preparation and guidance by the teacher. To be useful, there needs to be performance standards that describe expectations clearly enough that students and teachers evaluating the same work are likely to arrive at the same score. Students' self-assessment should reflect evaluation against performance standards. Models and anchor papers become important tools.

Decisions about Student Performances

The cycle of plan-teach-assess-adjust repeats itself many times in the weeks and months that make up a school year. Periodically–sometimes in every class–teachers decide that it is appropriate to judge students' work. These judgments usually result in the assignment of a grade for a particular piece of work.

One aspect of assessment is determining what statements or grades best reflect the students' performance. How this is accomplished varies greatly. When an assessment strategy is appropriate for the type of learning target being assessed, the teacher is likely to have better quality evidence upon which to base a judgment. Teachers can make more explicit statements about what the student knows or how the student performs.

Whereas the ongoing classroom assessment that influences

instruction and monitors students' progress is formative and often blends with instruction, summative assessment judges where the student is at a particular point in time in relation to set performance criteria. The results are recorded and are likely to carry different consequences than formative assessments. All assessment involves judgment, and judgment in most classrooms becomes a grade. Grading practices reflect local policies and vary from school system to school system. In fact, in most schools, the grades may represent different performance criteria from teacher to teacher. This is true for grades on individual samples of students' work and on report cards.

Judgements must be made against clear, public criteria so that everyone understands how the students will be graded.

Grades on report cards represent judgment about the work of students over an extended period of time. These may vary in what they represent even more than on individual pieces of work since teachers add the additional layer of aggregating grades from the entire grading period.

The grade of "C" in one class does not necessarily represent the same performance as a "C" in another class. Some teachers include grades for behavior and responsibility with their content grades. Other teachers "throw out" the lowest score during each grading period. Some teachers feel the mode is the best reflection of students' performance while other teachers use the mean. A score of 90 in one class is an "A-" on the scale of 90-100 = A. If the scale is 93-100=A, then the score of 90 becomes a "B." Even in primary grades where some schools choose to provide narrative evaluations, rather than numerical grades, the judgments of one teacher are not always congruent with those of other teachers about similar performance.

What is clear about grades and report cards is that they send powerful messages and have influence on how students feel about themselves as learners. For this reason, judgments must be made against clear, public criteria. Judgments that are summative should be based on the broadest, most valid information possible.

Decisions about programs

Teachers examine their programs in light of information they have from assessments. They want to know if they have strong

programs and how they might improve. Classroom assessment data are often supplemented by data from norm-referenced tests that the school system or the state requires such as the End-of Grade or End-of-Course tests. Many content areas have national standards which are resources that discuss program evaluation in greater detail.

Conclusions

Teachers have, or can get, information about students' performances that is rarely available to others. Since decisions and actions related to teaching and learning are most effective when based on accurate information, teachers should assume an active role in designing and carrying out quality assessments.

Some of the most important decisions teachers make are decisions about instruction. These decisions often begin before students come into the class and continue throughout the school year. Cumulative folders give general information about the students.

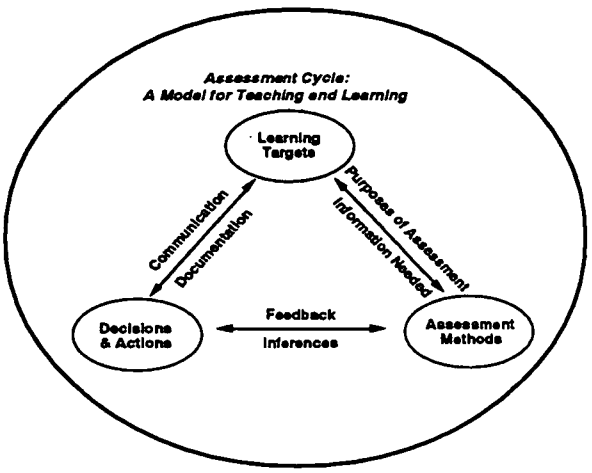
Once classes begin, teachers want to build upon what students know. Up-front assessment yields diagnostic information that helps the teacher know about the class as a whole and students as individuals. Teachers identify strengths and misconceptions, interests and concerns.

Through on-going classroom assessment teachers have a powerful tool to gather the varied evidence about students' learning that they need to make instructional decisions during the year. When students' thinking and reasoning can be examined, teachers are able to plan lessons that continue to promote achievement. Teaching matches the needs of the students; it goes beyond being a teacher activity carried out without regard to where students are.

During a school year a great deal of evidence about the progress students make is available to teachers. This assessment data is used to assign grades, but it is perhaps most powerful when it is used to make important decisions about teaching and learning.

Chapter 5

Documenting and Communicating



Before teachers decide upon the next set of learning targets for students, they want to document accurately and communicate clearly what students have achieved. While the process of communication and documentation is infused into all stages of the cycle of classroom assessment, it is particularly evident when teachers are ending one cycle of learning and beginning another.

The more data teachers have about students, the more they need an efficient way to manage that information.

Documenting student learning, which can be as simple as taking notes or keeping checklists, needs to be organized, systematic, and on-going. It is a wise investment of time to decide in advance what information to record and how to record it.

Communication needs to be carefully considered and thoughtfully planned. Like documentation, it needs to be on-going and reflective. It can range from impromptu to formal. It can include communication of expectations, of progress, and/or of achievement. Finally, communication is vitally interactive—not only does the teacher communicate to the student and parents, but students can also communicate what and how they have learned through student-led conferences.

Recording Assessment Information

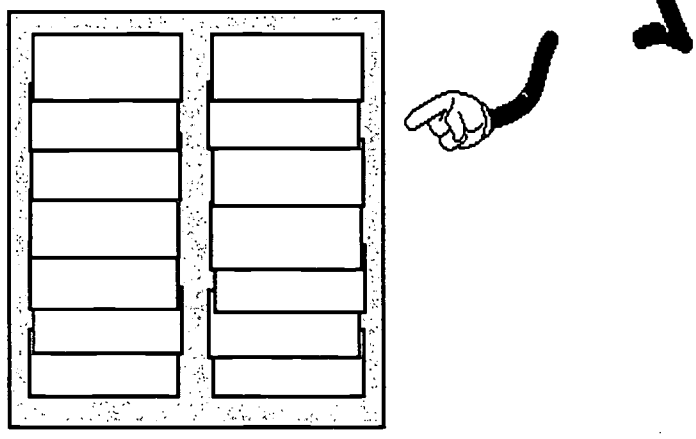
By using classroom assessment as a means of helping to make appropriate instructional decisions and to promote student achievement, teachers find that they know more about their students than they ever have. Teaching the class as a whole may (or may not) be their main instructional strategy, but they are able to fine tune lessons because of what they have learned about individuals.

For conversations and interviews to be helpful, teachers start with good questions and then become good listeners.

However, there is a downside to all of this. The more they know, the more teachers recognize the differences in the ways that students think and solve problems. For most teachers this increases the pressures they place on themselves to meet the needs of their students and to capitalize on students' interests and strengths. Accompanying this desire to individualize within the class as a whole is the need to keep records about groups and individual students. Almost always the dilemma arises:

How do you have time to make notes on all that you are learning about each student?

Attach index cards on a clipboard; tape across the top of the cards, attaching the bottom cards first.



File cards in students' folders as they are filled. Write names on bottom of cards so they are easy to see.

When observations and conversations are the primary means of assessment, the need to make some kind of anecdotal records is especially compelling. While there are no quick and easy solutions, there are many record keeping strategies that teachers may use to record their observations. Since there is no "one right way" to do this, the trick is for teachers to develop some systematic approach that is workable for themselves.

Teachers need to try several strategies as a means to develop their own system for recording, organizing, and summarizing the information they gather on their students. Becoming comfortable with making notes during class in some way that does not necessitate recopying the information but is easy to access and store usually takes some experimentation

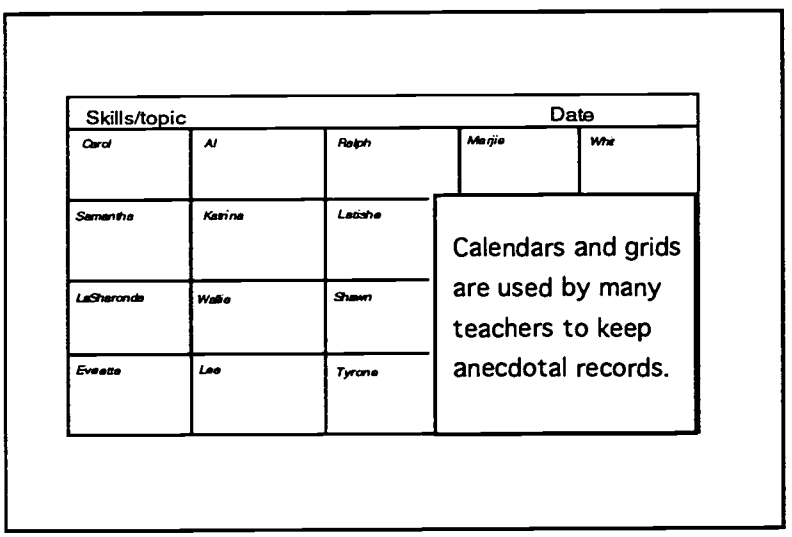
Some teachers create schedules for themselves so that they are focusing on specific students each day. These rotations can be by working groups, reading groups, alphabetical order (making notes about a certain number of students each day), or seating. Other teachers make notes about individuals as needed or during seatwork, homework, or finish-up times. This latter method usually generates many notes about some students and only a few, if any, about others.

There are many ways to use grids and calendars to record assessment information.

- A grid per objective(s)

or

- A grid for the week
- List every student
- Make notes on students for each topic or record observations only as needed



Notecards are a popular place to record information about students because they can be filed in students' folders for future reference. These are easily attached to a key ring or taped on a clipboard. They can be stored in "recipe boxes" and five or

six cards pulled each day for the teacher to write about the students. Different cards would be chosen the next day.

Grids with each student's name in a section, stick-on mailing labels, and sticky notes are other quick ways to record information. The grids have the advantage of allowing teachers to glance at information about the class as a whole related to a learning target. They can be filed into a notebook or folder for future reference. They have the disadvantage of not having all information about students in individual folders.

Mailing labels can be preprinted or teachers can write students' names on them as they make notes. Since they, like sticky notes, can be quickly stuck into students' folders, they have advantages of already being together for each individual at conference times.

There are many other record keeping strategies. Some teachers keep limited notes about students in their plan books; others use different checklists. Writing notes to students on their papers gives students feedback, but has the disadvantage of no record for the teacher for future reference unless the papers are placed into working folders or portfolios. Electronic notepads and computer-generated profile sheets are other ways to establish anecdotal records and document what teachers are observing throughout instructional sequences. Since these data may influence grades, dating entries is important.

Deciding What To Record

Not everything that goes on in the classroom can or should be written down. Answering questions of *why* are these notes necessary and *what* should be written down often will also answer questions of *how* is the best way to do this. Here are some suggestions that may be helpful in getting started with making anecdotal records.

- *Decide what you want to know.* Do you have this information about some students already? Do you need to assess everyone?
- *Decide if the information is relevant to your instructional planning.* Must this assessment result in a grade (evaluation) or is your purpose to make decisions about what to do next (or first)?

Teachers Share...

Record Keeping Ideas

Notes written directly on work samples or on dittoed forms for **Portfolios.**

Checklists, labels, and Electronic Records

- class roster or group rosters down left-hand side or page with skills, information, concepts, or processes to be assessed across the top
- skills clustered as taught or listed for a grading period
- mailing labels preprinted with students' names; added to individual folders
- computer programs which track student responses
- electronic notepads

Student Reflections

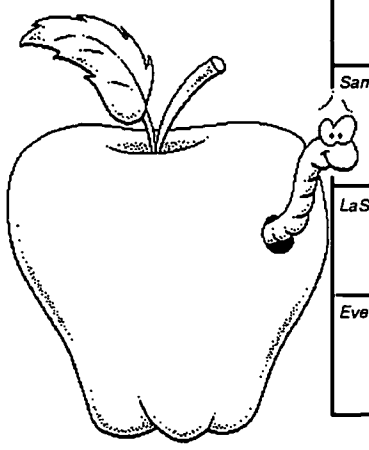
- student journals
- explanations of strategies
- responses to open-ended questions
- worksheets and drafts with student's annotations and corrections

Note Cards and Plan Book

- cards kept on a clipboard or in a file
- cards kept on students working in a group or on a specific project
- note cards completed on a rotating schedule
- note cards completed as needed - usually on students having special difficulty
- notes written during finish-up times
- notations in lesson plan book about lessons or about specific students
- notes in plan book about future groupings for help with specific skills

Matrix or Profiles

- summaries and synthesis of work over a period of time
- use anecdotal notes, observations, interviews, student products to evaluate student work
- highlight or mark to summarize performance at the end of grading periods
- summarizes on-going assessments



Skills/topic			Date	
Carol	*Al*	*Ralph*	*Marjie*	*Whit*
Samantha	*Katrina*	*Latishe*		
LaSharonda	*Wallie*	*Shawn*		
Everette	*Lee*	*Tyrone*		

Calendar or Grid

- grid per objective or cluster of objectives or
- grid for each week
- every student listed
- record observations as needed or
- use to make notes on all students about a specific topic
- for on-going assessment

- *Decide what are the simplest and most straight forward ways to attain and record the information you want.* Do you need to make comments or will a checklist do?
- *Decide if there is written work you can use.* Can student products and other student-made recordings make this task easier?
- *Assess what you value.* Are you assessing what is most important for students to know and looking for evidence of learning in a variety of situations?
- *Be specific in your own mind about what you want to assess.* Are you assessing a variety of learning targets?

Teachers who have developed record keeping strategies have suggestions about ways to begin:

- Focus on a few children at a time.
- Focus on a few objectives at a time.
- Keep comments clear, precise, and related to students' demonstrated understanding. Write down only what you see!
- Date your observations.
- Do not be too quick to give up on a method, but do not stick with it if you find it does not work!
- Find a fellow teacher with whom you can share the needs and successes of the students.
- Be quick to make notes of unexpected information, but be slow to decide that students have made new content their own.
- Look for evidence that students have learned in a variety of situations.
- Find a method that fits you so that you do not have to recopy notes. Make quick notes during class rather than trusting that you will find time later.

Rubrics as Tools for Communication

Rubrics are extremely useful in clarifying for students the teacher's expectations for quality work and in communicating to students an evaluation of their own work. Rubrics can also be important in clarifying learning targets: in fact, asking students to participate in the construction of a rubric to evaluate their work helps them to identify with clarity, precision, and accuracy what excellence is.

Rubrics are tools which list criteria for the desired accomplishment; they are descriptions of performance expectations linked to a scale for evaluating the work. They help differentiate among work that ranges from poor to excellent. While students may be the primary audience for rubrics, they save teachers time by describing "what counts" and clarifying how work will be evaluated. They lead to consistency in grading.

There are a number of ways to describe different kinds of rubrics. These descriptions give four basic types:

- Generic rubrics – broad expectations/criteria that cut across specific tasks and are helpful in writing more specific criteria for individual assignments,
- Holistic rubrics - descriptions which are based on the overall performance/work,
- Analytic rubrics - scores which are based on specific dimensions of the task, and
- Developmental rubrics- descriptions that are locations on a continuum of progress.

Knowing basic differences in the design of rubrics is helpful in deciding how teachers might begin to use them in their classrooms. Scales with three or more performance levels (score points) are usually used in all types of rubrics.

Generic rubrics define broad expectations at each performance level. They can apply to many different tasks. These general rubrics are helpful starting points for class discussions about what is "good work." They are also useful as references when specific rubrics are developed for a project or performance task. They may describe general attributes such as "clear, appropriate, and accurate" that become specific to individual assignments in analytic scoring.

Generic rubrics give general guidelines for scoring. *Holistic rubrics* evaluate the performance as a whole (the overall impression), while *analytic rubrics* score different attributes such as persuasion, appropriateness, accuracy, and process separately.

Although *holistic rubrics* can be based on separate criteria, they combine those criteria to evaluate performance as a whole. Holistic rubrics, which thus give one score, are well suited to tasks that can be performed or evaluated as a whole. They are also suitable for tasks for which students may not need extensive feedback.

Analytic rubrics assess specific dimensions of the task. It is much like judging a cake at the county fair. Criteria describe

how the product should smell, taste, and look. When the cake is finished, it can be rated based upon the expectations.

In *developmental rubrics* students' knowledge and skills are placed on a continuum of progress. Characteristics of readers at emerging, developing, and independent levels are an example. Students' writing is often examined using developmental rubrics that describe experimental, early, conventional, and independent levels.

Rubrics are more than a checklist. They help teachers evaluate complex performances where there may be more than one correct way of demonstrating learning.

Different rubrics also have different labels for the score points (performance levels). Numbers, words, and phrases identify the ranges. For example, some rubrics use scales of 0 to 2 or "Not Yet" to "Consistently." Other rubrics are labeled "Keep Trying" to "Wow," Levels I to IV, or "Novice" to "Expert." With 6 to 8 point scales there is greater precision but the rubrics are more difficult to create. Teachers must judge if the gains in precision are worth the extra time and effort. No matter what the scale or the labels, all rubrics use descriptors to distinguish between different levels of performance.

For example, "On time for class; not tardy" can be easily noted on a checklist, while "Demonstrates self-direction" may be demonstrated in many ways that teachers are likely to want to note.

Rubrics serve several purposes. The rubrics state teachers' specific expectations, define levels of proficiency, and establish specific goals for students (individually and collectively). They help teachers monitor student progress, assess the effectiveness of their instruction, and make summative statements about students' growth and development.

While rubrics are frequently used as scoring guides to evaluate students' work, they also help students understand the learning targets and what constitutes quality work. Using a rubric, students are able to self-monitor and improve their work as they complete assignments. Parents too better understand expectations, performance levels, and growth over time through rubrics.

At first writing rubrics is very hard. (Using them, by contrast, seems easy!) Teachers and students can work together to describe the expected "good work." Looking at different examples of rubrics used by others is helpful in this process. Comparing models of works that have been identified as quality and those which are judged as poor is also helpful. The next task is to list the characteristics of the quality

Holistic Rubric for Comparison/Contrast

4	A.. **Accuracy:** Examples selected by the student reflect complete accuracy with respect to the character, event, etc. B. **Well-Supported**: The student is consistently able to give evidence from one or more sources to explain or defend his examples. C. **Uniqueness**: The student gives several examples that are uncommon and show an original point of view. D. **Depth**: The student gives many examples that go beyond the obvious or use the obvious to gain insight to think beyond.
3	A. **Accuracy**: Examples selected by the student reflect nearly complete accuracy with respect to the character, event, etc. B. **Well-Supported**: The student is often able to give evidence from one or more sources to defend or explain his examples. C. **Uniqueness**: The student gives a few examples that are uncommon and show an original point of view. D. **Depth**: The student gives a few examples that go beyond the obvious or use the obvious to gain insight to think beyond.
2	A. **Accuracy:** Examples selected by the student reflect some accuracy with respect to the character, event, etc. B. **Well-Supported**: The student is sometimes able to give evidence from one or more sources to explain or defend his examples. C. **Uniqueness**: The student gives one or two examples that are uncommon and show an original point of view. D. **Depth**: The student gives one or two examples that go beyond the obvious or use the obvious to gain insight to think beyond.
1	A. **Accuracy:** Examples selected by the student reflect little or no accuracy with respect to the character, event, etc. B. **Well-Supported**: The student is unable to give evidence from one or more sources to explain or defend his examples. C. **Uniqueness**: The student gives no examples that are uncommon and show an original point of view. D. **Depth**: The student gives no examples that go beyond the obvious. *developed by teachers from Eau Claire, Wisconsin*

performance and of the poor performance. Intermediate descriptors identify performance that is beyond minimal but yet not sufficient. As rubrics are used with student work, they should be refined.

Writing rubrics may best be started while working with a colleague or with the students rather than the teacher working alone. The discussions will help clarify criteria and engage teachers and students in partnerships.

Portfolios

A portfolio is a purposeful collection of students' work which the teacher or the students have judged to be important evidence of what the students have learned. Portfolios might be compared to gradebooks in that they are the records of students' progress. Gradebooks house evidence that has been turned into letters or numbers. Portfolios, on the other hand, house the original evidence (student work) along with students' reflections about the work. Most portfolios are evaluated holistically, though sometimes pieces are individually scored.

Portfolios: In recent years portfolios have been a major part of most discussions of assessment. Used by every discipline, they are systematic collections of students' work designed to meet specific criteria. Depending upon their purpose, the guidelines vary for what goes into the portfolios, who is responsible for putting entries into them, and how they are evaluated. Portfolios are frequently used to encourage students' self-assessment and often are used in three-way conferences with families, students, and teachers. They may be housed in folders or pizza boxes or be designed as sophisticated electronic versions.

Portfolios offer students who seem to be disadvantaged by traditional testing additional opportunities to demonstrate what they have learned. Implemented without careful planning, however, they may become little more than daily work folders.

Portfolios are created for many different purposes:

- Working portfolios (folders) contain daily work and may be drawn from to create other portfolios.
- There are showcase (celebration) portfolios in which students' best work is displayed.
- Growth (chronological) portfolios are used to document progress over time.
- Certifying (documentation) portfolios may provide evidence of a student's employability or that the student has met criteria for a certain license. These are usually more structured in their contents.
- Anthologies (Wiggins, 1998) can be portfolios which contain a combination of on-demand tasks, best work, and samples to show growth over time.

The pieces that are included in the portfolios may be chosen by the student, by the teacher, or by both.

Portfolios are frequently used in a formative manner; they are put together, refined, and discussed periodically throughout the school year. Because portfolios contain direct evidence of students' achievements, and because in most situations students can take responsibility for selecting entries and culling their work, portfolios encourage students to play a key role in "telling their own stories." As students choose works that meet established guidelines, they look for quality performances and often recognize their own mistakes and areas that need improving. They reflect on the work they choose, explaining why each piece is important. Because portfolios can contain evidence in a variety of formats, teachers, families, and students are able to have a more in-depth view of the student's achievements than the "snapshot" that single assessments provide.

Advantages of Portfolios

There are obvious advantages and disadvantages that need to be considered when deciding to use portfolios. An important advantage has already been mentioned: portfolios allow students who have difficulty demonstrating all that they know through traditional tests to present their knowledge in other ways. This makes an important assumption, however, that students have ownership of the portfolio and choose to include significant selections. Poor choices by students can be named as a disadvantage of portfolios.

Some advocates of portfolios see them as more important for improving student learning than for documenting performance since portfolios require reflections and self-assessment. Because pieces can be chosen by students, teachers, or by both, students become decision-makers and are encouraged to focus on their strengths and to set goals to address their needs. In the process of reflecting on the pieces they include in their portfolios, students often pay more careful attention to criteria for quality work and can identify ways to improve on future assignments.

As teachers use portfolios for assessing the development of student learning, they have an in-depth view of students' achievements. Rather than comparing an individual's work with that of other students, portfolios focus on the work and progress of each student. Comparisons are made with students' previous work and/or against performance standards. Families better understand goals and have direct evidence of learning. Because portfolios may be designed to contain evidence from multiple sources including student work samples, on-demand tasks, teacher's observations, and traditional tests, they are helpful in making decisions about students' programs and placements.

Another advantage is that serious attention to portfolios is likely to influence the nature of assignments teachers make as they recognize the need to plan assessments that match learning targets and reveal students' thinking and reasoning, not just their recall. Exercises in which students underline subjects and verbs will indicate if students can identify these parts of speech in the context of the given sentences, but they do not indicate if students can write clear sentences with subject-verb agreement. Likewise, a worksheet filled with number facts does not tell much about how the student got answers (recall, counting objects, using fact strategies, or counting on fingers). The worksheet can be judged for right or wrong answers, but there is nothing to indicate how readily the student can use facts in solving problems.

Portfolios, say advocates, are likely to promote ongoing assessment that is closely integrated with classroom activities. Instruction and assessment blend as students refine their portfolio entries using the teacher's feedback, their own self-assessments, and peer reviews.

Disadvantages of Portfolios

Several disadvantages of portfolios are obvious. If teachers and students are not clear on their purpose and the guidelines for what will go into the portfolios, they are likely to become fat work folders for some students and additional ingredients of messy desks for others. Implementation of portfolios also requires professional development so that teachers can see samples of what works for others, decide the purposes and guidelines for their own class, discuss how they will use the portfolios, and learn how to score them.

Teachers must find the time to evaluate the portfolios – reflecting on the various learning targets the work samples represent and the contents as a whole. Scoring of portfolios may be difficult because of

- the variety of entries and often
- the absence of specific rubrics for classroom assignments or criteria for the portfolio as a whole.

Time, like teacher expertise, is a concern related to the using portfolios. Class time must be allocated for students to work with their portfolios and for teachers to conference with students. Since portfolios are most powerful when shared with families (increasingly often through student-led parent conferences), teachers must allocate time to prepare for sharing with families.

Take Time to Reflect

- What is my purpose in using portfolios?
- How frequently do I want students to work with them? When will they do this?
- How much choice will students have in determining what goes into the portfolio?
- How will we evaluate the portfolios? Will the portfolios relate to report cards? If so, how?
- What will happen to the portfolios at the end of the year?

Student-led Parent Conferences

Helping students assume responsibility for their own learning is important for citizenship and work in the 21st century. It is also important for motivation and achievement in school. Students are most likely to take ownership when they believe they can succeed. Their beliefs are likely to be more affirming when they understand the importance of the content, recognize the criteria for quality performances, and have opportunities to share their understandings with others. The school-to-work process begins at an early age when students develop skills in planning and organizing, communicating information, and responding to questions.

> One third grade teacher creates "Great Walls" of student work (often with rubrics) for the conferences. Since every student's work for the same assignments is displayed, families see how their student's work compares with others. They still are able to celebrate the individual student's progress yet have an understanding of the class standards.

Student-led parent conferences are an excellent vehicle for giving students responsibility and control. Students' achievement is validated when they present work samples to their parents, talk about their struggles and successes, and explain the criteria for determining good work. Confidence is built as students display and describe their work and explain their progress. The pride comes from jobs well-done and reinforces students' commitment to learning. Ownership and celebration go hand in hand, and students tend to set higher goals for themselves when they help to set the goals initially.

While powerful and important, student-led parent conferences are not designed to be a substitute for all teacher-student conferences or teacher-parent conferences. Greater responsibility for students does not diminish teacher responsibility.

Teachers remain the primary "planners" and as such should proceed with care when instituting student-led parent conferences. For example, student-led conferences should never occur without careful planning and coaching. Poorly executed conferences may lessen parents' confidence in the teacher and the school and destroy students' confidence rather than bolster it.

Success in Student-led Conferences

- Define clear goals for the conference. Be sure that students and families know these goals. Have the administration's support.
- Guide the collection and organization of data (student work) so that they are consistent with the goals of the conferences and with the learning targets that will be discussed.
- Give students opportunities to rehearse. All students need to be well-prepared to explain what they have accomplished and what they would like to do in the future.

Preparing for Student-led Parent Conferences

In order for students to successfully lead conferences, teachers must begin planning early. Students need to experience an instructional program that nurtures their ability to set goals and to reflect on and assess their learning. Informing students and families at the first of the year that they will be involved in student-led parent conferences helps to set the stage. By knowing the kinds of things they will be sharing with their families, students have a reason for keeping their work folders up to date and for paying special attention to rubrics the class uses throughout the grading period.

Having the support of the principal is critical. School administrators are most likely to be supportive of student-led parent conferences when they are involved from the beginning.

They quickly see the benefits when students are involved in self-assessment and setting goals. They also recognize the pitfalls if students are not well-prepared for the conferences. If several teachers work together to plan for student-led parent conferences, even more good ideas are likely to arise.

Students need an outline for what to do during the conference. Reflection sheets will help them summarize their work. These may be developed jointly with the students or teachers may choose to prepare them ahead of time. Teachers may also want some specific items (i.e., tests, problem solving assignments) to be included in the conference.

One middle school faculty in Raleigh has a three-conference program for sharing information with families. The initial conference at the end of the first grading period is between teachers and parents. Students are given copies of their report cards the week before the conferences and work in their homerooms to write a letter to their families about the grades. They are encouraged to provide any evidence that they would like for their teachers to share with the parents and to include in the letter a goal for each class during the next grading period. Report cards and the letters are given out by the teachers at the conferences. (Teachers confer with each other so that each teacher meets only with the families of the homeroom students.)

The second conference, held at the end of the second grading period (end of the semester), involves teachers, students, and families. The teachers and students share the responsibility for leading the discussion. In the week prior to the conferences several extended homeroom periods are devoted to preparation for the conferences. Again, students have copies of their grades; they must prepare an evaluation of how they did on the goals they had set at the end of first quarter and create statements about their goals during the next grading period.

The last scheduled conference is at the end of the third quarter with the same participants. Students are responsible for conducting the entire conference and spend time collecting more extensive evidence to explain the grades on their report cards as well as to discuss themselves as students. They are able to talk with all of their teachers prior to the conferences and receive specific assistance from their homeroom teacher. During the conference, the role of the teacher is to be a listener and to add clarification if necessary. Again, evaluating and setting goals are important aspects of the discussion.

Timelines and specific responsibilities for the conference process will vary according to students' ages and experiences. Letters of invitation, practice conferences with peers and teacher, and organizing materials need to take place in the days prior to the conferences. A table of contents for the work folders (portfolios) is helpful, especially if entries are a combination of teacher-selected and student-chosen work.

Asking families to write letters back to their student helps to bring teachers, students, and families together in supporting the student's achievement. Students may write thank you letters to parents for attending.

Some teachers choose to have three or four students conferencing with their families during the same time. Teachers move from group to group during the time allotted for the conference (usually 20 - 30 minutes). Students need to be clear on how and when to involve the teacher. Conference schedules should make provisions for teachers to join each group for a few minutes at the end of the conference, and teachers need to be clear with families that additional, private conferences may always be scheduled for a future date.

There is much more to be said about student-led parent conferences. Perhaps the best way to summarize the power of these conferences to examine comments by those involved.

Students benefit from leading conferences with their parents:

- They must save, organize, and select samples of work to share with their parents.
- They gradually recognize ways in which they need to be responsible for their work.
- They develop confidence and self-esteem and are better able to communicate with their parents.
- They engage in self-assessment and reflection about their work.
- They are motivated to set high, but appropriate goals for themselves.
- They are able to share their accomplishments and progress.
- They have opportunities to work with their peers through practicing.
- They have opportunities to use their oral language skills in purposeful ways.

Parents learn much from student-led conferences:

- They have opportunities for purposeful communication with their child.
- They perceive their child as a responsible learner.
- They release an appropriate amount of control for learning to their child.
- They understand the student's command of curriculum content and process.
- They appreciate the value of monitoring and measuring progress.
- They learn content expectations and criteria for determining successful learning.
- They compare their child's achievement with class standards and indirectly with work of other students.
- They recognize and appreciate the nature and quality of the classroom learning environment.
- They understand and appreciate the need for teachers to serve as facilitators.

Teachers benefit from student-led conferences:

- Maintaining a risk-free environment is easier as students develop confidence.
- Opportunities to observe parents and students working together give insight into the support students receive.
- Nurturing lifelong learners is possible when knowledge is important to families.
- The quality of learning escalates when children have criteria for monitoring and evaluating their own progress.
- Students recognize and assume responsibility for learning when they must share it with others.
- Shared responsibility and goals, and stronger, positive relationships are built.
- The facilitator role of the teacher is valued and understood in addition to the more traditional view of that teacher as an instructor.

Senior Projects

There are many ways that senior projects can be designed and implemented. However, all senior projects consist of three phases. In the first phase, students research and write a paper on an approved topic of their choice. This paper should demonstrate the acquisition of knowledge through researching, writing, interviewing, and applying complex skills such as problem solving.

The second phase includes the development of a project or product that is a reasonable and appropriate extension of the research paper. Students most often work with an adult or mentor who will provide guidance and who will verify the students' work.

Finally, the students deliver an oral presentation to a panel of judges, usually community members. This formal presentation is followed by a question and answer period. Students are usually required to provide a portfolio of the project components, including a resume.

For each component, students are assessed using a standardized scoring guide. The program is demanding for both teachers and students. Senior projects require rigor and commitment from students to complete satisfactorily this type of long-term assignment and to demonstrate complex knowledge and skills. These skills may include:

- gathering information through researching and reading,
- communicating information by writing, speaking, and using visual exhibits and verbal and nonverbal expression,
- using numbers, graphs, charts, drawings, and problem-solving techniques gained from math and science, and
- using current systems of technology.

There are numerous benefits for students, faculty, and the community at large. Senior projects are an enjoyable, culminating activity where students demonstrate what they have learned during grades K-12. They can bring collaboration and cohesiveness to high school faculties. Finally, they provide a public forum for students to demonstrate to the

community, to their teachers, and to each other what they know and what they can do.

The Importance of Student Responsibility and Self-Assessment

One important result of learning is for students to understand what they have learned, its importance, and the criteria for evaluating quality so well that they can assess their own work. This begins in the earliest years of schooling. Teachers want their students to assess themselves, and self-correct, while they engage in their learning or work, rather than waiting for the teacher to pass judgment. An independent, lifelong learner continuously self-assesses and continuously grows.

Self-assessment can also play a crucial role in motivating students. Students, even very young students, can learn to clarify learning targets for themselves, to establish criteria for quality performance, to improve their work as they monitor it, and to evaluate their growth over time. These actions can lead to feelings of independence, autonomy, and pride. As students hone their understanding of how they are progressing toward high levels of achievement, they not only become more successful learners, they also become more self-confident. Allowing students to judge their work against established performance standards, rather than comparing students' work with that of others, allows students to note their progress while recognizing what they need to improve upon.

Strategies for student responsibility

Students need encouragement to begin reflecting upon their thinking and learning. Often they need assistance in getting organized and developing power over their own learning. Assessment should be used to build confidence, not destroy it. Strategies which help students include:

- Provide a folder for students to store work and teach them how to file materials;
- Have students date all work;
- Model using criteria to complete assignments and rubrics to evaluate work;

- Make students responsible for identifying evidence of learning;
- Meet with students to discuss their work;
- Help students set clear goals and ask them to critique their work, telling good features as well as things that need improving;
- Make simple scoring sheets for students to use in evaluating their work;
- Provide opportunities for students to share their problem solving strategies;
- Take pictures of students and their work and display them prominently;
- Construct rubrics with students;
- Institute student-led parent conferences;
- Use reflective prompts for journals; and
- Celebrate improvements as well as attainment of goals

The Thorny Issue of Grades

Any chapter on documentation and communication must include a discussion of grades, since grades have been the most traditional and powerful way of documenting students' success and communicating it to students and parents. A teacher's grading system is different, however, from other forms of documentation and communication. For example, a portfolio is a collection of the student's original work; observation notes are descriptions of what students are doing. Grades, however, are symbols of the student's achievement and as such they can be ambiguous. It goes without saying that different teachers grade in different ways. It is equally true that the teacher may grade differently in different situations. So what do those grades, or symbols, represent?

Most teachers traditionally have graded their students using three methods:

1. norm -referenced assessments, where students' work is compared to other students' work, usually in a "bell-shaped curve"
2. criteria- referenced assessment, where student's work is compared to standards of excellence
3. growth models, where students' work is compared to their own past performances.

While many teachers have abandoned the practice of norm-referenced grades, they still struggle with combining criteria -referenced assessments with growth models. On one hand, teachers understand that all students need to reach high levels of achievement; criteria -referenced assessments help students understand how closely they approached these targets. On the other hand, teachers question the fairness of not considering students' growth, especially since they enter the class with such diverse levels of achievement.

This problem is just one reason that the subject of grading in complex and thorny. Other difficult questions teacher ask themselves include the following:

1. Should effort ever count? How should it be treated? If a student is allowed multiple attempts to master a learning target, how many attempts are reasonable?

2. Have I clearly defined, for myself and for my students, what an A means in this class?

3. Does that grade of A really reflect my priorities? Are the important learning targets the actual basis of individual grades? Does the way I put separate grades together for a collective grade skew the priorities?

4. How should I put grades together? Should I count early grades which may indicate initial efforts when the student is still learning? What will happen if I count only grades that come later in the learning unit?

Rick Stiggins, in *Student-Centered Classroom Assessment*, offers the following advice about grading. While it does not answer all questions, it does provide a useful overview.

1. Begin with the subject matter knowledge that students are to master. Outline or write the actual information in order of importance to capture the most significant knowledge targets.

2. Decide what reasoning and problem-solving proficiency students must demonstrate; again, begin specific and note priorities.

3. Outline the skills, if any, the students must master or be able to do. List these skills and any key elements in successful performance.

4. Decide what products, if any, students are to create. Outline the nature of the products and the key elements of a high-quality product.

5. What dispositions are important and how will you know when students demonstrate them?

Although Stiggins does not recommend grading for attitudes, you may wish to consider important dispositions such as the ability to consider different perspectives or fair-mindedness, perseverance, and flexibility.

According to Stiggins, this planning should be the basis of grades, since grades should reflect the most important learning targets of your class. Turning this "big picture" into a concrete grading plan requires that you decide how you will assess each category and allocate weight to those categories.

Take Time to Reflect:

Stiggins also offers the following advice:

- Grade on achievement of prespecified targets only, not intelligence, effort, attitude, or personality.
- Always rely on the most current information available about student achievement.
- Devise grades that reflect achievement status with respect to preset targets rather than improvement.
- Decide borderline cases with additional information on achievement.
- Keep grading procedures separate from punishment.
- Change all policies that lead to miscommunication about achievement.
- Add further detail to grade reports when needed.
- Expect individual accountability for learning even in cooperative environments.
- Give extra credit for evidence of extra learning--not just for doing extra work!

Do I agree with these statements? Why or why not?

Consider a faculty discussion of these statements. How much agreement would I expect to find among my colleagues? What benefit may such a discussion have for our school?

References

Bright, G. and J. Joyner (1998). *Classroom Assessment in Mathematics: Views from a National Science Foundation Working Conference.* Lanham, MD: University Press of America.

Mager, R.F. (1962). *Preparing Instructional Objectives.* Palo Alto, CA: Fearon Publishers.

National Council of Teachers of Mathematics. (1995). *Assessment Standards for School Mathematics.* Reston, VA: National Council of Teachers of Mathematics.

Richardson, K. (1997). *Math Time: The Learning Environment.* Norman, OK: Educational Enrichment.

Stiggins, R.J. (1997). *Student-centered Classroom Assessment* (2nd ed.). Upper Saddle River, NJ: Merrill.

Wiggins, G. (1997). CLASS: *North Carolina State Assessment (1998* ed.). Princeton, NJ: the Center on Learning, Assessment, and School Structure.

Glossary

ability	The capacity to do something; the power to perform.
access	Rights and means to approach or engage in with understanding.
accountability	Proof of the educational impact of education. The concept of trying to hold appropriate parties accountable for their performance; in education these are usually administrators, teachers and/or students.
achievement	Successful accomplishment or attainment of educational goals.
achievement levels	Achievement levels allow for the comparison of student and group performance to established standards based on what is expected in a subject area at each grade level. In North Carolina, these predetermined levels of achievement relate the judgement of North Carolina teachers to the actual student performance on North Carolina tests.
achievement test	A test that measures the extent to which a person has "achieved" something, acquired certain information, or mastered certain standardized sets of skills - usually as a result of planned instruction or training.
alternative assessment	Generally, any kind of assessment technique other than traditional norm-referenced or criterion-referenced pencil-and-paper tests.
analytic scoring	A type of rubric scoring which separates the whole into the main components, traits, or characteristics of what is being assessed.

aptitude test	A test designed to measure a student's innate ability in a given area. Tests are available to measure aptitudes for many specific school subjects such as foreign language, art, music, and mathematics.
assessment	The process of gathering evidence; describing and analyzing information about reasoning, skills, knowledge, and performance. The term "assess" comes from Latin, meaning "to sit beside."
authentic assessment	Assessment which involve activities that are faithful, comprehensive representations of the contexts and complexity found in important, real-life performances of adults that are nonroutine yet meaningful and engaging for students.
benchmark	The measurement of group performance at various developmental levels against an established standard.
bias	Conditions in content, procedures, or interpretation of assessment information that favor one or more groups of participants over other groups.
checklist	A written list of performance criteria associated with a particular activity or product in which an observer marks the student's performance on each criterion according to a scale that only has two choices (observed, not observed)
coherence	The quality of logical connection and orderly relationship of parts to each other and to the whole.
cohort	A group whose progress is followed by means of measurements at different points in time.
competency test	A test intended to make sure a student has met certain minimum or essential standards of skills and knowledge.
concepts	General and fundamental ideas.

consistency	Compatibility or agreement among successive acts, ideas, or events.
construct	A theoretical idea developed to describe and to organize some aspect of existing knowledge.
content standards	Standards that specify "what students should know and be able to do" in various subjects and domains, such as mathematics or applied learning. Content standards detail the knowledge, skills, and other necessary understandings that school should teach in order for all students to attain high levels of competency in the subject matter. In North Carolina, the content standards are described in the *Standard Course of Study* for each subject, course, and/or grade,
context	The circumstances or situation in which a problem occurs or which skills and strategies can be applied.
credibility	The quality of being plausible, believable, dependable, or worthy of confidence.
criteria	A statement of the most valued characteristics of a performance.
criterion-referenced test	A test designed to provide information on the specific knowledge or skills possessed by a student. Such tests usually cover relatively small units and are related to instruction. Performance is measured in reference to the mastery of particular skills. Scores from these tests have meaning in terms of *what* the student knows or can do, rather than in relation to the scores made by some external reference (or norm) group.
curriculum alignment	The degree to which a curriculum's scope and sequence matches the instruction (material taught) and the evaluations measures (assessments used).

curriculum-based test	A test that is designed to measure student performance in reference to goals and competencies specifically defined in a curriculum. Curriculum-based tests are a specific set of criterion-referenced tests and have a similar purpose. Curriculum-based tests may also be referred to as domain-referenced tests.
cut score	Performance score used to determine the passing of minimal competency levels.
developmental scale scores	Developmental scales are similar to rulers that allow growth in a subject to be measured over time. Developmental scale scores are intended to be interpreted much like height in inches. As is the case for growth in height, we expect an annual increase in scale scores, but not the same increase every year. The rate of growth is a bit faster in the earlier grades than in the later grades.
diagnostic test	A test that is used to diagnose, analyze, or locate a student's specific areas of weakness and/or strength to determine the nature of the weakness or deficiencies, and, where possible, to suggest the cause. Diagnostic tests are most commonly prepared for the skill areas.
disposition	Interest in, and appreciation for a tendency to think and act in positive ways; includes confidence, curiosity, perseverance, flexibility, inventiveness, and reflectiveness.
distractor	An incorrect choice in a multiple-choice or matching item (also called a foil).
equal	Having the same quantity, measure, value, privileges, status or rights.
equitable assessment	The degree to which the process of gathering evidence has provided opportunities equally appropriate for each student to demonstrate the valued thinking processes, knowledge, and skills that he or she has developed.

equivalent	Equal in value or meaning, interchangeable, or having comparable effects.
essay test	A response format that requires students to answer questions in writing, emphasizing evaluation and elaboration rather than choosing a correct alternative. The responses could be brief or extensive.
evaluation	The process of determining the worth of, or assigning a value to, something on the basis of careful examination and judgement.
formative assessments	On-going assessments used to monitor student growth and progress. They provide feedback but not final evaluation.
framework	An organizing system for, and arrangement of understanding, performances, and dispositions to be assessed, which will assist the planning of assessments.
generalization	An inference or conclusion based on evidence and supported by a theory describing the relationships between the evidence and the inference or conclusion.
goal	The expected end results of education stating what schools are expected to accomplish and what students are expected to learn.
high-stakes testing	Any testing program whose results have important consequences for students, teachers, schools, and/or school districts. Such stakes may include promotion, certification, rewards/ sanctions, graduation or denial/approval of services or opportunities.
holistic scoring	A type of rubric scoring based on the belief that what is being assessed is more than the sum of its parts and can be either specifically linked to a written criteria (*focused holistic rubric*) or implied (*general impression rubric*) most frequently occurring data point.

inferences	Conclusions or assertions derived from evidence; deductions.
IQ (intelligence quotient) tests	The first of the standardized norm-referenced tests, developed more than a century ago, designed to measure a person's native intelligence. Used more today for psychological screening purposes.
item	An individual question or exercise in a test or evaluative instrument.
item analysis	The process of analyzing each item on a test to determine proportions of students selecting each answer. Can be used to diagnose particular strengths and weaknesses of students, as well as the test's validity or possible bias. *Difficulty.* The percent of students who answer an item correctly, designated as *p*. *Discrimination.* The extent to which an item differentiates between high-scoring and low-scoring students.
judgements	Authoritative estimates or opinions of quality, value, and other features, formed by distinguishing the relations among multiple sources of sound and reasonable evidence; formal decisions.
learning targets	What students are expected to understand and be able to do; for example, the goals and objectives in the North Carolina *Standard Course of Study*
mean	A measure of central tendency that is computed by summing all of the data points and then dividing by the number of data points. The mean is one of several ways of representing a group with a single, typical score. It is figured by summing all of the scores in a group and dividing by the number of people in the group.
measure	To indicate how much of some specified, quantifiable unit is present; to assign numbers to variations in a quantifiable attribute or trait. For

	conditions that cannot be quantified with sufficient certainty and accuracy for the intended purposes, a description is more appropriate than measurement.
measurement-based assessment	Procedures for (1) developing and selecting test items or assessment tasks by the degree to which they differentiate among examinees and for (2) administering and scoring such tests to provide statistically adequate scores for making valid generalizations regarding the psychological traits, attributes, and skills that the test items or tasks are designed to measure.
median	A measure of central tendency. The point on the scale which separates the group into two equal subgroups; one way of representing a group of scores with a single typical score. Unlike the mean, the median is not affected by extremely low or high scores.
metacognition	The knowledge individuals have of their own thinking processes and strategies, and their ability to monitor and regulate these processes. This requires learners to analyze, reflect on and monitor one's own learning. Metacognition, i.e., knowledge, awareness and control of cognition, is an outcome of conscious reflection.
mode	A measure of central tendency. The mode names the most frequently occurring data points.
multiple-choice item	A response format in which the student is asked to choose from a list of possible options the one correct *or* the one best response to the given question.
norm	A performance standard that is established by a reference group and that describes average or typical performance. Usually norms are determined by testing a representative group of students and then calculating the group's test performance.

norm-referenced test	A test that is designed to measure performance in relation to a comparison group. A norm-referenced test tells how the scores of each student or group of students compares to the scores of the original group that took the test.
normal distribution	The scores of students do not necessarily produce the same distribution as the scores of the norm group. A distribution of scores or other measures that in graphic form has a distinctive bell-shaped appearance. In a normal distribution, the measures are distributed symmetrically about the mean. Cases are concentrated near the mean and decrease in frequency, according to a precise mathematical equation, the farther one departs from the mean.
normal curve equivalent (NCE)	A normal curve equivalent (NCE) is a normalized student score with a mean of 50 and a standard deviation of 21.06. NCEs range from 1 to 99. They allow comparison between different tests for the same student or group of students and between different students on the same test. NCEs have many of the same characteristics as percentile ranks, but have the additional advantage of being based on an interval scale. That is, the difference between two successive scores on the scale has the same meaning throughout the scale. NCEs are often required by many categorical funding agencies (for example, Title I).
objective	A stated, desirable outcome of education that represents a portion of a larger goal.
observation	A systematic way to collect data by watching or listening to students during an activity.
open-ended item	A response format in which the student is asked to create a written response, where the correct response may vary - there is not simply one correct answer or there is more than one strategy for arriving at the answer, for example, an essay item. The score scale of the response depends on the

	justification, rationale, or explanation that supports the response. A higher score on the item is dependent on answering all parts of the item, rather than only responding to part of the item correctly. Open-ended items engage students in situations and allow students at many levels of understanding to begin working on the problem, to make their own assumptions, to develop creative responses, and to effectively communicate their solutions
opportunity to learn	The degree to which a student has been exposed to the learning experiences needed to meet high academic standards, which is largely a function of the capacity and performance of the courses and schools the student has attended. Equitable opportunities to learn consist of equal chances for learning, with equally appropriate, favorable, or advantageous combinations of circumstances (i.e., opportunities to learn are equitable when they are responsive to the same extent to each student's needs).
outcome	Learning, results or consequences.
percentile	Percentiles range from a low of 1 to a high of 99, with 50 denoting average performance. The percentile rank corresponding to a given score indicates the percentage of a reference group obtaining scores equal to or less than that score. For example, if a student scores at the 65th percentile, it means that he or she performed better on the test than 65% of the norm group. Caution: Students and parents should be reminded that the percentile rank does not refer to the percentage of items answered correctly. Real differences in performance are greater at the ends of the percentile range than in the middle.
performance assessment	A response format in which the student is asked to apply knowledge and skills, actively; an assessment task that requires the student to create an answer or product to demonstrate his or her

	knowledge or skills; for example, a debate or a science investigation. A score on the task is dependent on not only the final product, but also the process that is undertaken to produce the product.
performance standards	Standards that answer the question "How good is good enough?" Performance standards establish the degree or quality of students' performance in the subject matter set out by the content standards. Performance standards are often illustrated by exemplars of student work and may contain gradations of performance, i.e., needs improvement, meets standard, above standard.
portfolio	A file of student work that is accumulated over time and is centered on a particular topic or content area. Portfolios are an assessment technique where specified contents are reviewed according to established criteria to determine the level of the student's performance and progress.
process	A generalizable method of doing something, generally involving a number of steps or operations which are usually ordered or interdependent.
program evaluation	The process of determining the effectiveness of an educational program in achieving its goals and, therefore, its value in comparison to its required resources.
project	A complex assignment that expects more than one type of activity and production for completion. A type of performance assessment.
quality	Degree of excellence. The quality of assessment evidence is characterized primarily by the authenticity of the tasks, the reliability of the sample of evidence, and the credibility of the evidence for the intended purposes.
quartile	One of three points that divide the data points in a distribution into four groups of equal size. The first quartile, or 25th percentile, separates the

	lowest fourth of the group; the middle quartile, the 50th percentile or median, divides the second fourth of the cases from the third; and the third quartile, or 75th percentile, separates the top quarter.
range	A measure of variability. The range is the difference between the largest (highest) and the smallest (lowest) data points.
raw score	The number of test items a student answers correctly.
reliability	The measure of consistency for assessment instruments. A reliable test will yield similar scores when abilities or knowledge are similar across time.
representativeness	The degree to which assessment evidence represents the valued thinking processes, knowledge, and skills that the student has developed.
rubric	A set of ranked criteria to give direction to the scoring of assessment tasks or activities. To be useful, a scoring rubric must be derived from careful analysis of existing performances of varying quality and expectations. A **task-specific rubric** describes levels of performance for a particular complex performance task and guides the scoring of that task consistent with relevant performance standards. A **general rubric** is an guideline for creating task-specific rubrics, or for guiding expert judgement, where task-specific scoring rules are internal to the scorer.
sampling	A way to get information about a large group by examining only a small number of the group (*the sample*), or by giving all members small segments of the test. When conducted properly, the results are considered highly reliable and generalizable.

scale scores	Scores that express how far a given raw score is from a reference point, usually the mean score. Scale scores are units on an equal interval scale. The difference between two successive scores on the scale has the same meaning throughout the scale. The advantage of scale scores is that they allow an individual's performance to be compared on several different tests, regardless of the raw score scales.
scoring	The process of discriminating among performances according to differing levels of quality and assigning a descriptive label or number to the performance. In **holistic scoring,** the entire performance as a whole is considered, and one label or number is assigned. In **analytical scoring,** separate scores are assigned to fundamentally different dimensions of the performance.
self-assessment	Personal skills which provide information from ones proximal and past actions so as to identify personal options for use with potential goals to improving individual learning. The framework or purpose for the act of self monitoring is understood and applied.
self-evaluation	Applying a personal or an external framework on the ones proximal and past actions so as to make a decision about what course of action to follow.
self-reflection	Personally analyzing and making judgements about what has happened but not necessarily with an internal or external framework for analysis. May include only a description of what has happened.
skill	Ability to perform routine procedures. In mathematics, procedures typically consist of computational or manipulatory methods. In reading, procedures typically consist of semantic, syntactic, or graphaphonic decoding.

standard

A statement about what is valued that can be used for making a judgment of quality. The level, type and possibly format of a performance expected for a targeted group at a point in time or along a discrete or continuous skill. In assessment, a level of performance established by tests users of different scores obtained by students. The word standard is also used in curriculum development to denote a level of achievement or an outcome toward which all students should strive.

standard deviation

A measure of variability. The standard deviation is an indication of the amount of variation in a set of data points.

Usually two-thirds of data points fall within one standard deviation above or below the mean. When the standard deviation of a set of data points is small, most of the data points in the set are scattered far from the mean. When this number is large, most of the data points in the set are clustered close to the mean. In other words, students perform very differently from each other when the standard deviation is large and very similar to each other when the standard deviation is small.

standard error of measurement

As applied to a single test score, the standard error of measurement (sem) is the amount by which the score may differ from a hypothetical true score due to errors of measurement. The larger the sem, the less reliable the obtained test score. In about two-thirds of the test scores, the obtained score will not differ by more than one standard error of measurement from the true score.

stanine

Stanine scores are scores that have been converted to a standard scale to allow easy comparison across different test forms, classes, grade levels, schools and districts. Comparison is easy because the mean is set at 5, the standard deviation is set at approximately 2,

	and the range is set at 1 to 9. In general, stanines of 1 to 3 are considered below average, 4 to 6 are considered average, and 7 to 9 are above average. Stanines, like percentile ranks, indicate a student's relative standing in a norm group.
standardized test	A test designed to provide a systematic sample of individual performance, administered according to prescribed directions, scored in conformance with definite rules, and interpreted in reference to certain normative information. An assessment instrument given to a large number of persons under similar conditions, designed to yield comparable scores.
summative assessment	Assessments that are used at the end of an instructional sequence to provide an overall evaluation of student achievement.
task	A goal-directed assessment activity, demanding that students use their background knowledge and skills in a continuous way to solve a complex problem or question.
teacher-made test	A test prepared by the teacher for use usually in the classroom. It can be norm-referenced (designed to measure differences among the individuals in the class) or criterion-referenced (specifying minimum levels of acceptable performance), and performance-based. Since teacher-made tests are usually not standardized, the scores bear no relation to those obtained by students on a given test in other classrooms.
valid	Justifiable, well grounded, sound; producing the desired results, efficacious; incontestable.
valid inferences	Justifiable assertions and conclusions that lead to and support desirable results. Justification is made primarily on the quality of the evidence and its adequacy for the intended purposes and their consequences.

validity

The extent to which a test measures that which it purports to measure. All procedures for determining test validity are concerned with the relationship between performance on the test and other independently observable facts,

Content Validity. The extent to which the content of the test represents a balanced and adequate sampling of the outcomes (domain) about which inferences are to be made.

Criterion-Related Validity. The extent to which scores on the test are in agreement with (concurrent validity) or predict (predictive validity) some criterion measure.

Construct Validity. The extent to which a test measures some relatively abstract psychological trait or construct; applicable in evaluating the validity of tests that have been constructed on the basis of an analysis of the trait and its manifestation.

The validity of an assessment indicates the degree of accuracy of either predictions or inferences based upon the score from the assessment. The reasonableness, quality, and efficacy of an assessment for particular educational purposes, decisions, and consequences are important issues of validity.

variance

A measure of variability. The variance is derived by subtracting the mean of the set of data points from each data point, squaring the difference, summing the squares, and dividing by the number of data points.

If the variance of the set is small, then the data points are close together; if the variance is large, then the data points are more spread out.

TM

STUDY GUIDE

FOR

CLASSROOM ASSESSMENT: LINKING INSTRUCTION AND ASSESSMENT

Guide has been produced by
Instructional and Accountability Services
North Carolina Department of Public Instruction
and
SouthEastern Regional Vision forEducation

January 1999

About this Study Guide

This study guide is a companion to the manual *Classroom Assessment: Linking Instruction and Assessment*, published by the North Carolina Department of Public Instruction, January 1999. In order to complete this study guide, you will need that manual and access to the video that has been provided for each school. It is also recommended that you find a colleague or a small group with whom to work, although you may complete the study guide independently. This study guide assumes that you have participated in a three-hour overview provided by your school administration.

Some of the benefits for choosing to work through this study guide include:

- You will gain a higher level of awareness of the role assessment plays in helping students learn.

- You will improve your own assessment capacity by building greater assessment literacy and by increasing your professional knowledge about how assessment is currently viewed.

- If your principal or central office administration provides a three-hour overview and establishes renewal credit for this study guide, you can receive 2.5 renewal credits for the 25 contact hours involved in completing the work in this study guide.

- You can include the work in this study guide as a way to meet the goals you set in your school improvement plan or your professional development plan.

A special thanks to Nancy McMunn, consultant with SERVE, who collaborated with Jeane Joyner and Jan Williamson in producing this document and to Dee Brewer, Carolyn Cobb, Jane Cowan, Julie Malone, Sherron Pfeiffer, and Rhonda Welfare for reviewing this document.

If you have questions or comments about the classroom assessment manual or this study guide, please contact Jeane Joyner in Accountability Services (phone: 919-715-1864 or email jjoyner@dpi.state.nc.us) or Jan Williamson in Instructional Services (phone: 919-715-1875 or email jwilliamson@dpi.state.nc.us).

TABLE OF CONTENTS

Before You Begin this Study Guide

In this study guide you will be responsible for completing the following:

Assigned Readings: (Each individual completes)

- ❑ Complete the assigned readings for each section of this guide. The readings will be found in the manual *Classroom Assessment: Linking Instruction and Assessment* or the specific article will be provided.

Questions for discussion and reflection: (Complete with a peer or team)

- ❑ This section provides you with insightful questions that will help begin dialogue on specific assessment issues pertinent to the section of the manual. You are encouraged to form peer teams or dialogue groups at your school to sit and share/discuss your responses and thoughts about these issues before you proceed to other activities.

Activities: (Can be completed individually or in a team)

- ❑ You will be given several activities in this part; some will be required and others are optional.

Checks for understanding: (Each individual completes)

- ❑ This section provides you with an opportunity to assess your understanding of the information in this module. You are asked to complete this part by responding to the assessment included in each section and reviewing any information you do not fully understand.

Reflections on what was learned: (Each individual completes)

- ❑ This section asks you to reflect on what you have learned in reference to the goals of the study guide. When you have completed the work in the entire study guide, you will be asked to respond in writing to some of these questions.

The Role of Assessment in Teaching and Learning

Overview

Goals:

The focus of this chapter is to review the information you were presented in the introductory workshop for this assessment module. As you work through this chapter you will:

1. Reflect on the chapter and dialogue with your peers to understand that assessment goes beyond state and local testing.
2. Begin the process of building assessment literacy and assessment capacity for taking action based on your own reflections.

Estimated time for this module: **3 hours**

Assigned readings:

Overview chapter in the manual: pages 1-18

Questions for discussion and reflection:

- When you are learning something new what helps you learn it?
- What do you think "teaching to the test" means to you or your peers?
- What are the assessment principles that guide you in the classroom?
- What previous training have you receive in assessment?
- If someone asked you to explain what "quality" assessment means, what would you tell him or her?

Required activities:

- Activity 1 - Principles for Quality Classroom Assessment
- Activity 2 - Video Segment—Watch "Beyond Testing"

Optional activities (Choose one):

- Activity 3 - Shifts in Assessment Practices
- Activity 4 - Environments that Promote Learning

Checking for understanding:

Complete the true/false quiz.

Reflecting on what was learned:

1. How do the needs of assessment information vary for the state, district, school, teacher, parent, and student?

2. What questions or insights do I now have about assessment that I have not previously thought about?

Complete this checklist when you have finished this module:

_____ Read pp. 1-18 in the manual

_____ Discussed "Questions for discussion and reflection"

_____ Completed Activity 1 - "Principles of Quality Classroom Assessment"

_____ Completed Activity 2 - video segment "Beyond Testing"

_____ Completed one activity of my choice

_____ Completed "Checking for understanding"

_____ Reflected on what was learned

Key for Checking for Understanding (page 11 in this study guide):

1. False see page 2 of *Classroom Assessment: Linking Instruction and Assessment*
2. True see page 3 of *Classroom Assessment: Linking Instruction and Assessment*
3. True see page 7 of *Classroom Assessment: Linking Instruction and Assessment*
4. True see page 8 of *Classroom Assessment: Linking Instruction and Assessment*
5. False see page 8 of *Classroom Assessment: Linking Instruction and Assessment*
6. True see page 11 of *Classroom Assessment: Linking Instruction and Assessment*
7. False see page 13 of *Classroom Assessment: Linking Instruction and Assessment*
8. False see page 13 of *Classroom Assessment: Linking Instruction and Assessment*
9. True see page 14 of *Classroom Assessment: Linking Instruction and Assessment*
10. False see page 17 of *Classroom Assessment: Linking Instruction and Assessment*

ACTIVITY 1

PRINCIPLES FOR QUALITY CLASSROOM ASSESSMENT

Directions:
Read each of the following principles (pp. 16-18) and respond to either *a* **or** *b*

a) What does this mean for you and your students in your classroom?

or

b) What questions do you need answered in order to understand what this assessment principle means?

Principle	Personal Response
Principle 1: Quality assessments arise from and accurately reflect clearly specified, appropriate, and essential learning targets.	
Principle 2: Quality assessments are specifically designed and focused to serve instructional purposes.	
Principle 3: Quality assessments accurately reflect the intended target and serve the intended purpose.	
Principle 4: Quality assessments promote equity by providing ample opportunities for students to demonstrate their learning.	
Principle 5: Quality assessments provide sufficient evidence of learning to permit confident conclusions (inferences) about student achievement.	
Principle 6: Quality assessments are designed, developed and used in ways that eliminate bias or distortion that may interfere with the accuracy of results.	

ACTIVITY 2

VIDEO SEGMENT: BEYOND TESTING

Directions:

Watch the fourth video, "Beyond Testing," on the Annenberg videotape that has been provided to each school. This is a portion of a series on classroom assessment that was produced by WGBH Television in Boston for the Annenberg/CPB Math and Science Project.

At the end of the manual of *Classroom Assessment: Linking Instruction and Assessment* in a special appendix, guide pages are provided for the Annenberg video series. The guide pages for "Beyond Testing" (pages 13-25) give an overview of this video and suggested questions for discussion.

ACTIVITY 3

Shifts in Assessment Practices

Directions:
(Read each statement as "I...," and check the box that indicates your classroom practice. For example, If "I assess a broad range of learning targets," I would check box 5. If "I assess mainly students' knowledge of specific facts and skills," then I would check box 1. If I believe that I fall somewhere in between the two practices, I would check box 4, 3 or 2.)

I...	WHERE AM I					I ...
Assess a broad range of learning targets	5	4	3	2	1	Assess mainly students' knowledge of specific facts and skills
Compare students' performance with established criteria	5	4	3	2	1	Compare students' performance with that of other students
Create my own assessments and observations	5	4	3	2	1	Use only assessments that come with the manual/book
Make the assessment process public, participatory, and dynamic	5	4	3	2	1	Make the assessment process secret, exclusive, and fixed
Allow individual students to demonstrate their knowledge in different ways	5	4	3	2	1	Ask all students to demonstrate their knowledge in the same way
Involve students in creating assessments	5	4	3	2	1	Develop assessments by myself
Allow students to retake tests that they do poorly on	5	4	3	2	1	Give students a single opportunity to take a test
Treat most assessment as incorporated into instruction	5	4	3	2	1	Treat most assessment as separate from instruction
Base inferences on multiple sources of evidence	5	4	3	2	1	Base inferences on a few sources of evidence
View students as partners in the assessment process	5	4	3	2	1	View students as those we are responsible for evaluating
Regard assessment as a continual process	5	4	3	2	1	Regard assessment as primarily summative
Hold self, parents, and students as accountable for assessment results	5	4	3	2	1	Hold only students accountable for assessment results

Calculate your score:
1. A score ranging between 48 – 60 means that your assessment practices reflect the current shifts.
2. A score ranging between 24 – 48 means that your assessment practices represent a combination of the old and new beliefs.
3. A score ranging between 12 –24 means your assessment practices have not changed to reflect the current shifts in assessment.

So, what does this mean? This self-assessment should help guide you in knowing where on a continuum your assessment practices are. Developing a clearer understanding will help you reflect on "quality" assessment practices that will help students learn and take greater responsibility for their own learning.

ACTIVITY 4

Scenarios: Environments that Promote Learning

Directions: Read the scenarios provided and respond to the questions at the end.

You are a parent moving to a community in North Carolina. You are very concerned about this move because your two children are both in middle school. They are anxious about moving and having to change schools mid year. Prior to your move, you decide to visit two schools in the area in which you want to live. These visits will help you decide on the best location for buying a new house based on the quality of the school. Here is what you observed during your visits.

Scenario One

FULCRUM MIDDLE SCHOOL (850 STUDENTS)

The principal greets you upon arrival at the school and immediately gives you a tour of the school. He points out to you what he feels are the important features of the school. He tells you about each feature as you walk through the building. Some of these special features are

- a huge technology laboratory with special technology teachers on staff prior to school hours, during school hours and after school hours,
- a library resource lab with Internet connections for 5 or more students at a time,
- a newly renovated area that allows for large groups of children to be outdoors at any given time,
- a renovated cafeteria that allows students four choices for meals,
- a new faculty lounge where teachers have several new copy machines, fax machines and telephone connections,
- bulletins boards in the halls next to each teacher room for displays,
- a huge dance and art studio for students, and
- a video recording room with all the latest technology. (No students or teachers were present and the principal told you he was in the process of hiring a part-time person to maintain the room.)

Once the tour is over the principal takes you to his office and asks you if you have any questions about his wonderful school. You think for a minute and ask how the school performs on statewide tests. The principal tells you that, based on end-of-grade tests, the school is in the top 20% statewide and everyone is working to improve this record. However, he says that this is not the most important issue for this school because the kids always do well on tests. The principal receives a call and tells you that if you think of other questions to call him later. He smiles and dismisses you in order to respond to the caller. You leave with no further questions.

SCENARIO TWO

CRISTAE MIDDLE SCHOOL (700 STUDENTS)

Two students greet you at the door of the school, announcing that they are your tour guides. They say that if you have any questions they cannot answer they will find someone who can. They have their tour and schedule planned out. You look around for adults and do not see any so you go with the students.

First stop: The Technology Lab

It is clean and neat but has some outdated computers. The students say they spend at least an hour twice a month in this lab to update their portfolios; they show you one of their electronic portfolios that houses the work they do in all their classes. You asked them what they liked about this process. They say their portfolios let them know what they are learning and what they need to improve.

Second stop: The P.E. Field

The students show you two areas. One area is a mile-long running track that the PE teachers use to help students work on their cardiovascular condition or that students use for walking. Another area is the football and soccer field that is used for exercise during the school day and for athletic games after school hours. Walking back inside, the students point out the patio area with benches and plants which the science club had planted.

Third Stop: The Cafeteria

The students warn you that it will be noisy since lunch has already begun. You notice teachers at several tables, sitting and talking to the students. The noise level is not as loud as you expect, and you notice that students clean up their mess. Your tour guides tell you about a student in-house committee at the school to help maintain a clean and safe environment. This group sets the rules for cafeteria behavior and also decides on any punishments.

Fourth Stop: The Faculty Lounge

The tour guides explain that this area is off-limits to students. When you look inside you see teacher mail boxes, a message center, and several computers hooked to the Internet here. There is only one old copy machine and one fax. The students say that there are telephones in every classroom.

Fifth Stop: Several classrooms

Your guides choose several classrooms to visit. Each teacher greets you and continues with the lesson. You are able to view several more student portfolios and ask the students questions. You spend a long time in the classrooms. When you ask one teacher about student performance on state and district tests, she shows you an action plan that teachers used to set goals to work toward improvement. She explains that this school was once ranked as low performing by the state but has made exemplary growth in the last few years.

Last Stop: The Principal's Office

The assistant principal gives you a packet of information about the school and goes over the contents. He asks if you have any further questions.

WHAT TYPES OF ENVIRONMENTS FOR LEARNING ARE REPRESENTED?

Based on the information you collected during your visits, which school would you chose for your children? Why?

If you could revisit either school, what else would you ask or like to see?

Checking for Understanding

Selected Response Test for Overview

(Please respond to questions for a self-check of your understanding of the content in this overview. This self-assessment models a selected response format. Look up information to clarify your understanding. An answer key is on page 4. The important thing is not your score, but the information you gained and how this information affects your assessment practices.)

True	False	Question
		1. Assessment literacy does not include knowing how to use assessment to motivate and teach students.
		2. The purpose of the assessment guides the types of information that needs to be gathered.
		3. No single type of assessment can meet all of the purposes of assessment or information needs of the various educational decision-makers.
		4. There are many other ways to communicate student achievement besides using numbers.
		5. The key assessment users are teachers.
		6. Assessment methods should match the learning targets.
		7. When making promotion/retention decisions, state test information is all the evidence for student performance that is needed.
		8. The kinds of inferences about student learning do not differ across methods of assessment.
		9. The "Model for Teaching and Learning," representing the assessment cycle is applicable for state, school, and classroom level decisions about assessment.
		10. Bias and distortion are not important when designing, developing and using assessments.

Classroom Assessment and Instruction

Chapter One

Goals:

The focus of this section is to gain understanding of assessment as a process. As you finish this section you will:

1. Reflect on a model for teaching and learning that involves an understanding of the assessment cycle.
2. Gain an understanding of the major components involved in the assessment process (feedback, purpose, targets, methods, uses, actions).
3. Understand how changing assessments practices can impact student achievement.

Estimated time for this section: 2 hours

Assigned readings:

Chapter 1 in the manual, "Classroom Assessment and Instruction," pages 19-34

Questions for discussion and reflection:

1. How does thinking of assessment as a process help teachers meet the needs of individual students within the context of the class as a whole?
2. If you had to describe what classroom assessment should reflect to a new teacher or a parent – what would that vision be?
3. How can thinking of assessment as a process for providing feedback positively impact student achievement?
4. How is assessment as a process different from assessment as an event that teachers do at the end of instruction?
5. Think of an instructional unit taught in your classroom that describes the instruction and assessment cycle noted on page 19 of the manual. How would the process be applied in your classroom? What problems or concerns do you have with this process?

Required activity:

- Activity 1- Read the article "Inside the Black Box" that follows page 13 of this study guide and answer the accompanying question.

Checking for understanding:

Part I: Fill in the blank graphic for the assessment cycle.
Part II: Define the terms used to relate assessment as a process.

Reflecting on what has been learned:

1. How does thinking of assessment as an on-going process in the classroom support and enhance state testing?

2. Give examples to illustrate how assessment as a process relates to different purposes for assessment.

Complete this checklist when you have finished this module:

_____ Read pp. 19-34 in the manual

_____ Completed Activity 1 – "Inside the Black Box"

_____ Discussed "Questions for discussion and reflection"

_____ Completed "Checking for Understanding"

_____ Reflected on what was learned

ACTIVITY 1

INSIDE THE BLACK BOX

Directions:

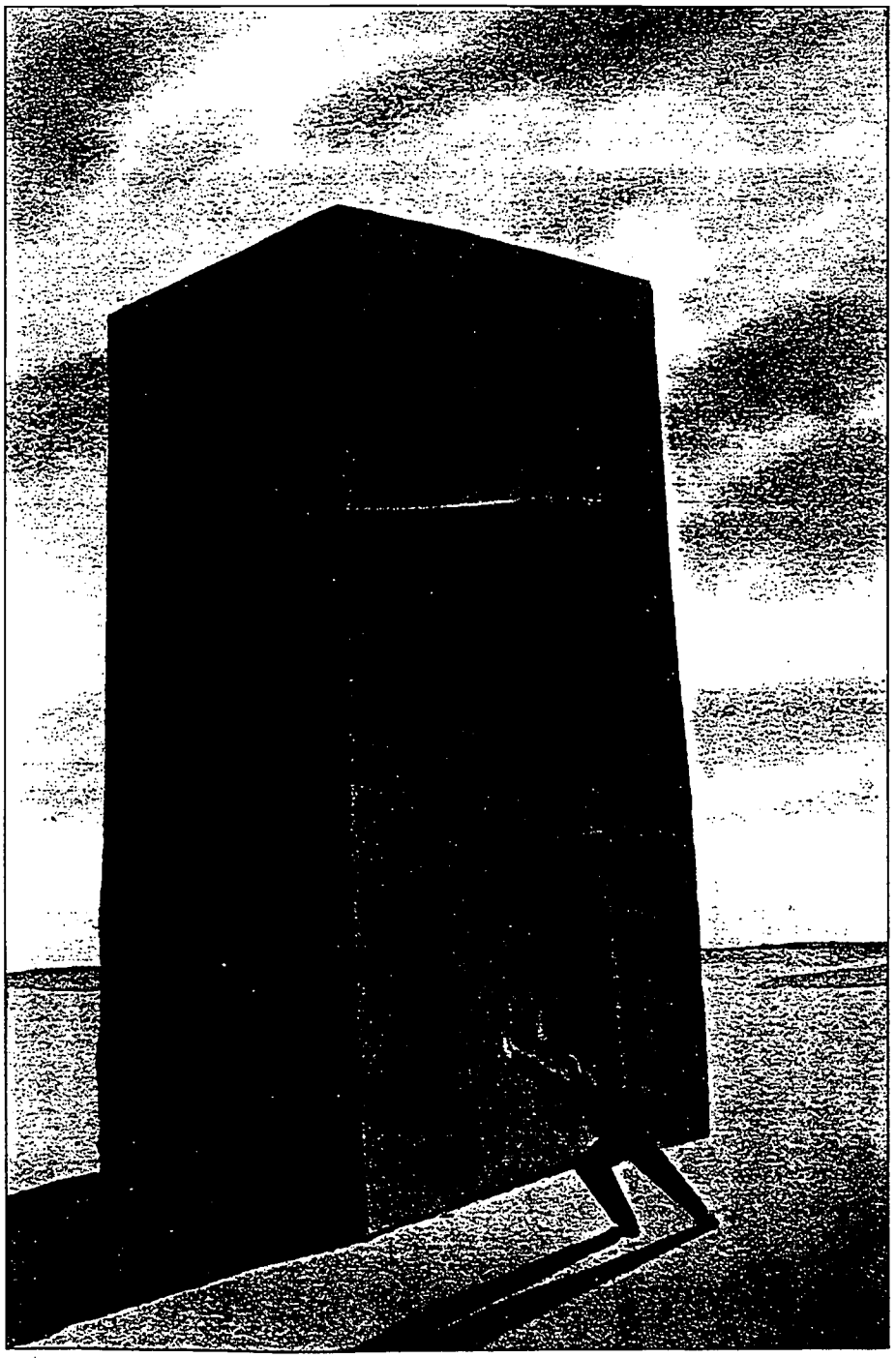
Read the article "Inside the Black Box," reprinted with permission from *Phi Delta Kappan*, October 1998. Respond to the following question:

In what ways does this article reinforce or extend your personal observations of how on-going classroom assessment fosters student achievement?

Inside the Black Box

Raising Standards Through Classroom Assessment

BY PAUL BLACK AND
DYLAN WILIAM



Firm evidence shows that formative assessment is an essential component of classroom work and that its development can raise standards of achievement, Mr. Black and Mr. Wiliam point out. Indeed, they know of no other way of raising standards for which such a strong prima facie case can be made.

RAISING the standards of learning that are achieved through schooling is an important national priority. In recent years, governments throughout the world have been more and more vigorous in making changes in pursuit of this aim. National, state, and district standards; target setting; enhanced programs for the external testing of students' performance; surveys such as NAEP (National Assessment of Educational Progress) and TIMSS (Third International Mathematics and Science Study); initiatives to improve school plan-

PAUL BLACK is professor emeritus in the School of Education, King's College, London, where DYLAN WILIAM is head of school and professor of educational assessment.

Illustration by A. J. Garces

ning and management; and more frequent and thorough inspection are all means toward the same end. But the sum of all these reforms has not added up to an effective policy because something is missing.

Learning is driven by what teachers and pupils do in classrooms. Teachers have to manage complicated and demanding situations, channeling the personal, emotional, and social pressures of a group of 30 or more youngsters in order to help them learn immediately and become better learners in the future. Standards can be raised only if teachers can tackle this task more effectively. What is missing from the efforts alluded to above is any direct help with this task. This fact was recognized in the TIMSS video study: "A focus on standards and accountability that ignores the processes of teaching and learning in classrooms will not provide the direction that teachers need in their quest to improve."[1]

In terms of systems engineering, present policies in the U.S. and in many other countries seem to treat the classroom as a black box. Certain *inputs* from the outside — pupils, teachers, other resources, management rules and requirements, parental anxieties, standards, tests with high stakes, and so on — are fed into the box. Some *outputs* are supposed to follow: pupils who are more knowledgeable and competent, better test results, teachers who are reasonably satisfied, and so on. But what is happening inside the box? How can anyone be sure that a particular set of new inputs will produce better outputs if we don't at least study what happens inside? And why is it that most of the reform initiatives mentioned in the first paragraph are not aimed at giving direct help and support to the work of teachers in classrooms?

The answer usually given is that it is up to teachers: they have to make the inside work better. This answer is not good enough, for two reasons. First, it is at least possible that some changes in the inputs may be counterproductive and make it harder for teachers to raise standards. Second, it seems strange, even unfair, to leave the most difficult piece of the standards-raising puzzle entirely to teachers. If there are ways in which policy makers and others can give direct help and support to the everyday classroom task of achieving better learning, then surely these ways ought to be pursued vigorously.

This article is about the inside of the black box. We focus on one aspect of teaching: formative assessment. But we will show that this feature is at the heart of effective teaching.

The Argument

We start from the self-evident proposition that teaching and learning must be interactive. Teachers need to know about their pupils' progress and difficulties with learning so that they can adapt their own work to meet pupils' needs — needs that are often unpredictable and that vary from one pupil to another. Teachers can find out what they need to know in a variety of ways, including observation and discussion in the classroom and the reading of pupils' written work.

We use the general term *assessment* to refer to all those activities undertaken by teachers — and by their students in assessing themselves — that provide information to be used as feedback to modify teaching and learning activities. Such assessment becomes *formative assessment* when the evidence is actually used to adapt the teaching to meet student needs.[2]

There is nothing new about any of this. All teachers make assessments in every class they teach. But there are three important questions about this process that we seek to answer:

- Is there evidence that improving formative assessment raises standards?
- Is there evidence that there is room for improvement?
- Is there evidence about how to improve formative assessment?

In setting out to answer these questions, we have conducted an extensive survey of the research literature. We have checked through many books and through the past nine years' worth of issues of more than 160 journals, and we have studied earlier reviews of research. This process yielded about 580 articles or chapters to study. We prepared a lengthy review, using material from 250 of these sources, that has been published in a special issue of the journal *Assessment in Education*, together with comments on our work by leading educational experts from Australia, Switzerland, Hong Kong, Lesotho, and the U.S.[3]

The conclusion we have reached from our research review is that the answer to each of the three questions above is clearly yes. In the three main sections below, we outline the nature and force of the evidence that justifies this conclusion. However, because we are presenting a summary here, our text will appear strong on assertions and weak on the details of the justification. We maintain that these assertions are backed by evidence and that this backing is set out in full detail in the lengthy review on which this article is founded.

We believe that the three sections below establish a strong case that *governments, their agencies, school authorities, and the teaching profession should study very carefully whether they are seriously interested in raising standards in education.* However, we also acknowledge widespread evidence that fundamental change in education can be achieved only slowly — through programs of professional development that build on existing good practice. Thus we do not conclude that formative assessment is yet another "magic bullet" for education. The issues involved are too complex and too closely linked to both the difficulties of classroom practice and the beliefs that drive public policy. In a final section, we confront this complexity and try to sketch out a strategy for acting on our evidence.

Does Improving Formative Assessment Raise Standards?

A research review published in 1986, concentrating primarily on classroom assessment work for children with mild handicaps, surveyed a large number of innovations, from which 23 were selected.[4] Those chosen satisfied the condition that quantitative evidence of learning gains was obtained, both for those involved in the innovation and for a similar group not so involved. Since then, many more papers have been published describing similarly careful quantitative experiments. Our own review has selected at least 20 more studies. (The number depends on how rigorous a set of selection criteria are applied.) All these studies show that innovations that include strengthening the practice of formative assessment produce significant and often substantial learning gains. These studies range over age groups from 5-year-olds to university undergraduates, across several school subjects, and over several countries.

For research purposes, learning gains of this type are measured by comparing the average improvements in the test scores of pupils involved in an innovation with the range of scores that are found for typical groups of pupils on these same tests.

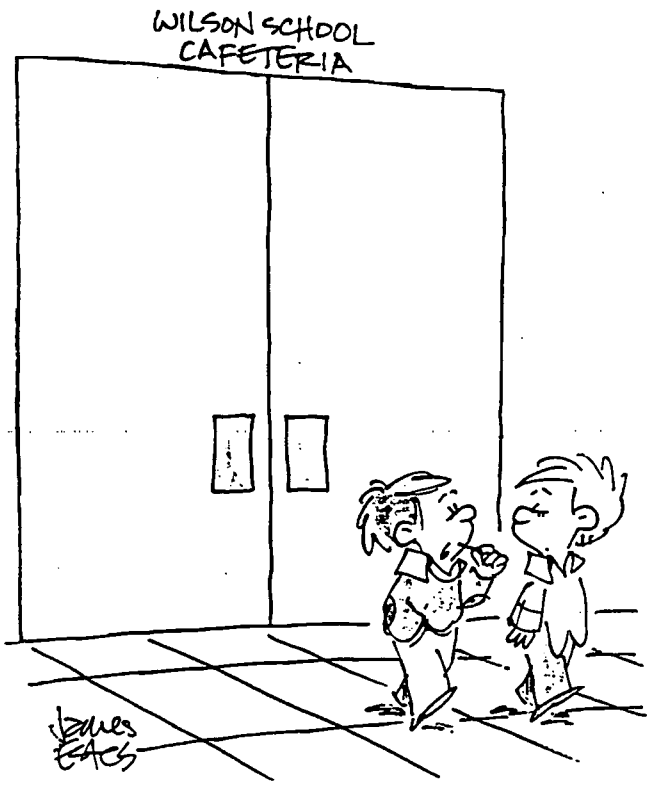
The ratio of the former divided by the latter is known as the *effect size*. Typical effect sizes of the formative assessment experiments were between 0.4 and 0.7. These effect sizes are larger than most of those found for educational interventions. The following examples illustrate some practical consequences of such large gains.

- An effect size of 0.4 would mean that the average pupil involved in an innovation would record the same achievement as a pupil in the top 35% of those not so involved.
- An effect size gain of 0.7 in the recent international comparative studies in mathematics[5] would have raised the score of a nation in the middle of the pack of 41 countries (e.g., the U.S.) to one of the top five.

Many of these studies arrive at another important conclusion: that improved formative assessment helps low achievers more than other students and so reduces the range of achievement while raising achievement overall. A notable recent example is a study devoted entirely to low-achieving students and students with learning disabilities, which shows that frequent assessment feedback helps both groups enhance their learning.[6] Many gains for such pupils could be particularly important. Furthermore, pupils who come to see themselves as unable to learn usually cease to take school seriously. Many become disruptive; others resort to truancy. Such young people are likely to be alienated from society and to become the sources and the victims of serious social problems.

Thus it seems clear that very significant learning gains lie within our grasp. The fact that such gains have been achieved by a variety of methods that have, as a common feature, enhanced formative assessment suggests that this feature accounts, at least in part, for the successes. However, it does not follow that it would be an easy matter to achieve such gains on a wide scale in normal classrooms. Many of the reports we have studied raise a number of other issues.

- All such work involves new ways to enhance feedback between those taught and the teacher, ways that will require significant changes in classroom practice.
- Underlying the various approaches are assumptions about what makes for effective learning — in particular the assumption that students have to be actively involved.
- For assessment to function formatively, the results have to be used to adjust teaching and learning; thus a significant aspect of any program will be the ways in which teachers make these adjustments.
- The ways in which assessment can affect the motivation and self-esteem of pupils and the benefits of engaging pupils in self-assessment deserve careful attention.



"The food's really not half bad, but the atmosphere leaves a lot to be desired."

Is There Room for Improvement?

A poverty of practice. There is a wealth of research evidence that the everyday practice of assessment in classrooms is beset with problems and shortcomings, as the following selected quotations indicate.

- "Marking is usually conscientious but often fails to offer guidance on how work can be improved. In a significant minority of cases, marking reinforces underachievement and underexpectation by being too generous or unfocused. Information about pupil performance received by the teacher is insufficiently used to inform subsequent work," according to a United Kingdom inspection report on secondary schools.[7]
- "Why is the extent and nature of formative assessment in science so impoverished?" asked a research study on secondary science teachers in the United Kingdom.[8]
- "Indeed they pay lip service to [formative assessment] but consider that its practice is unrealistic in the present educational context," reported a study of Canadian secondary teachers.[9]
- "The assessment practices outlined above are not common, even though these kinds of approaches are now widely promoted in the professional literature," according to a review of assessment practices in U.S. schools.[10]

The most important difficulties with assessment revolve around three issues. The first issue is *effective learning*.

- The tests used by teachers encourage rote and superficial learning even when teachers say they want to develop understanding; many teachers seem unaware of the inconsistency.
- The questions and other methods teachers use are not shared with other teachers in the same school, and they are not critically reviewed in relation to what they actually assess.
- For primary teachers particularly, there is a tendency to emphasize quantity and presentation of work and to neglect its

The ultimate user of assessment information that is elicited in order to improve learning is the pupil.

quality in relation to learning.

The second issue is *negative impact.*

• The giving of marks and the grading function are overemphasized, while the giving of useful advice and the learning function are underemphasized.

• Approaches are used in which pupils are compared with one another, the prime purpose of which seems to them to be competition rather than personal improvement; in consequence, assessment feedback teaches low-achieving pupils that they lack "ability," causing them to come to believe that they are not able to learn.

The third issue is the *managerial role* of assessments.

• Teachers' feedback to pupils seems to serve social and managerial functions, often at the expense of the learning function.

• Teachers are often able to predict pupils' results on external tests because their own tests imitate them, but at the same time teachers know too little about their pupils' learning needs.

• The collection of marks to fill in records is given higher priority than the analysis of pupils' work to discern learning needs; furthermore, some teachers pay no attention to the assessment records of their pupils' previous teachers.

Of course, not all these descriptions apply to all classrooms. Indeed, there are many schools and classrooms to which they do not apply at all. Nevertheless, these general conclusions have been drawn by researchers who have collected evidence — through observation, interviews, and questionnaires — from schools in several countries, including the U.S.

An empty commitment. The development of national assessment policy in England and Wales over the last decade illustrates the obstacles that stand in the way of developing policy support for formative assessment. The recommendations of a government task force in 1988[11] and all subsequent statements of government policy have emphasized the importance of formative assessment by teachers. However, the body charged with carrying out government policy on assessment had no strategy either to study or to develop the formative assessment of teachers and did no more than devote a tiny fraction of its resources to such work.[12] Most of the available resources and most of the public and political attention were focused on national external tests. While teachers' contributions to these "summative assessments" have been given some formal status, hardly any attention has been paid to their contributions through formative assessment. Moreover, the problems of the relationship between teachers' formative and summative roles have received no attention.

It is possible that many of the commitments were stated in the belief that formative assessment was not problematic, that it already happened all the time and needed no more than formal acknowledgment of its existence. However, it is also clear that the political commitment to external testing in order to promote competition had a central priority, while the commitment to formative assessment was marginal. As researchers the world over have found, high-stakes external tests always dominate teaching and assessment. However, they give teachers poor models for formative assessment because of their limited function of providing overall summaries of achievement rather than helpful diagnosis. Given this fact, it is hardly surprising that numerous research studies of the implementation of the education reforms in the United Kingdom have found that formative assessment is "seriously in need of development."[13] With hindsight, we can see that the failure to perceive the need for substantial support for formative assessment and to take responsibility for developing such support was a serious error.

In the U.S. similar pressures have been felt from political movements characterized by a distrust of teachers and a belief that external testing will, on its own, improve learning. Such fractured relationships between policy makers and the teaching profession are not inevitable — indeed, many countries with enviable educational achievements seem to manage well with policies that show greater respect and support for teachers. While the situation in the U.S. is far more diverse than that in England and Wales, the effects of high-stakes state-mandated testing are very similar to those of the external tests in the United Kingdom. Moreover, the traditional reliance on multiple-choice testing in the U.S. — not shared in the United Kingdom — has exacerbated the negative effects of such policies on the quality of classroom learning.

How Can We Improve Formative Assessment?

The self-esteem of pupils. A report of schools in Switzerland states that "a number of pupils . . . are content to 'get by.' . . . Every teacher who wants to practice formative assessment must reconstruct the teaching contracts so as to counteract the habits acquired by his pupils."[14]

The ultimate user of assessment information that is elicited in order to improve learning is the pupil. There are negative and positive aspects of this fact. The negative aspect is illustrated by the preceding quotation. When the classroom culture focuses on rewards, "gold stars," grades, or class ranking, then pupils look for ways to obtain the best marks rather than to improve their learning. One reported consequence is that, when they have any choice, pupils avoid difficult tasks. They also spend time and energy looking for clues to the "right answer." Indeed, many become reluctant to ask questions out of a fear of failure. Pupils who encounter difficulties are led to believe that they lack ability, and this belief leads them to attribute their difficulties to a defect in themselves about which they cannot do a great deal. Thus they avoid investing effort in learning that can lead only to disappointment, and they try to build up their self-esteem in other ways.

The positive aspect of students' being the primary users of the information gleaned from formative assessments is that negative outcomes — such as an obsessive focus on competition and the attendant fear of failure on the part of low achievers — are not inevitable. What is needed is a culture of success, backed by a belief that all pupils can achieve. In this regard, formative assessment can be a powerful weapon if it is communicated in the right way. While formative assessment can help all

pupils, it yields particularly good results with low achievers by concentrating on specific problems with their work and giving them a clear understanding of what is wrong and how to put it right. Pupils can accept and work with such messages, provided that they are not clouded by overtones about ability, competition, and comparison with others. In summary, the message can be stated as follows: *feedback to any pupil should be about the particular qualities of his or her work, with advice on what he or she can do to improve, and should avoid comparisons with other pupils.*

Self-assessment by pupils. Many successful innovations have developed self- and peer-assessment by pupils as ways of enhancing formative assessment, and such work has achieved some success with pupils from age 5 upward. This link of formative assessment to self-assessment is not an accident; indeed, it is inevitable.

To explain this last statement, we should first note that the main problem that those who are developing self-assessments encounter is not a problem of reliability and trustworthiness. Pupils are generally honest and reliable in assessing both themselves and one another; they can even be too hard on themselves. The main problem is that pupils can assess themselves only when they have a sufficiently clear picture of the targets that their learning is meant to attain. Surprisingly, and sadly, many pupils do not have such a picture, and they appear to have become accustomed to receiving classroom teaching as an arbitrary sequence of exercises with no overarching rationale. To overcome this pattern of passive reception requires hard and sustained work. When pupils do acquire such an overview, they then become more committed and more effective as learners. Moreover, their own assessments become an object of discussion with their teachers and with one another, and this discussion further promotes the reflection on one's own thinking that is essential to good learning.

Thus self-assessment by pupils, far from being a luxury, is in fact *an essential component of formative assessment.* When anyone is trying to learn, feedback about the effort has three elements: recognition of the *desired goal*, evidence about *present position*, and some understanding of a *way to close the gap* between the two.[15] All three must be understood to some degree by anyone before he or she can take action to improve learning.

Such an argument is consistent with more general ideas established by research into the way people learn. New understandings are not simply swallowed and stored in isolation; they have to be assimilated in relation to preexisting ideas. The new and the old may be inconsistent or even in conflict, and the disparities must be resolved by thoughtful actions on the part of the learner. Realizing that there are new goals for the learning is an essential part of this process of assimilation. Thus we conclude: *if formative assessment is to be productive, pupils should be trained in self-assessment so that they can understand the main purposes of their learning and thereby grasp what they need to do to achieve.*

Dialogue with the teacher provides the opportunity for the teacher to respond to and reorient a pupil's thinking.

The evolution of effective teaching. The research studies referred to above show very clearly that effective programs of formative assessment involve far more than the addition of a few observations and tests to an existing program. They require careful scrutiny of all the main components of a teaching plan. Indeed, it is clear that instruction and formative assessment are indivisible.

To begin at the beginning, the choice of tasks for classroom work and homework is important. Tasks have to be justified in terms of the learning aims that they serve, and they can work well only if opportunities for pupils to communicate their evolving understanding are built into the planning. Discussion, observation of activities, and marking of written work can all be used to provide those opportunities, but it is then important to look at or listen carefully to the talk, the writing, and the actions through which pupils develop and display the state of their understanding. Thus we maintain that *opportunities for pupils to express their understanding should be designed into any piece of teaching, for this will initiate the interaction through which formative assessment aids learning.*

Discussions in which pupils are led to talk about their understanding in their own ways are important aids to increasing knowledge and improving understanding. Dialogue with the teacher provides the opportunity for the teacher to respond to and reorient a pupil's thinking. However, there are clearly recorded examples of such discussions in which teachers have, quite unconsciously, responded in ways that would inhibit the future learning of a pupil. What the examples have in common is that the teacher is looking for a particular response and lacks the flexibility or the confidence to deal with the unexpected. So the teacher tries to direct the pupil toward giving the expected answer. In manipulating the dialogue in this way, the teacher seals off any unusual, often thoughtful but unorthodox, attempts by pupils to work out their own answers. Over time the pupils get the message: they are not required to think out their own answers. The object of the exercise is to work out — or guess — what answer the teacher expects to see or hear.

A particular feature of the talk between teacher and pupils is the asking of questions by the teacher. This natural and direct way of checking on learning is often unproductive. One common problem is that, following a question, teachers do not wait long enough to allow pupils to think out their answers. When a teacher answers his or her own question after only two or three seconds and when a minute of silence is not tolerable, there is no possibility that a pupil can think out what to say.

There are then two consequences. One is that, because the only questions that can produce answers in such a short time are questions of fact, these predominate. The other is that pupils don't even try to think out a response. Because they know that the answer, followed by another question, will come along in a few seconds, there is no point in trying. It is also generally the case that only a few pupils in a class answer the teacher's questions. The rest then leave it to these few, knowing that they cannot respond as quickly and being unwilling to risk making mistakes in public. So the teacher, by lowering the level

Tests given in class and tests and other exercises assigned for homework are also important means of promoting feedback.

of questions and by accepting answers from a few, can keep the lesson going but is actually out of touch with the understanding of most of the class. The question/answer dialogue becomes a ritual, one in which thoughtful involvement suffers.

There are several ways to break this particular cycle. They involve giving pupils time to respond; asking them to discuss their thinking in pairs or in small groups, so that a respondent is speaking on behalf of others; giving pupils a choice between different possible answers and asking them to vote on the options; asking all of them to write down an answer and then reading out a selected few; and so on. What is essential is that any dialogue should evoke thoughtful reflection in which all pupils can be encouraged to take part, for only then can the formative process start to work. In short, the dialogue between pupils and a teacher should be *thoughtful, reflective, focused to evoke and explore understanding, and conducted so that all pupils have an opportunity to think and to express their ideas.*

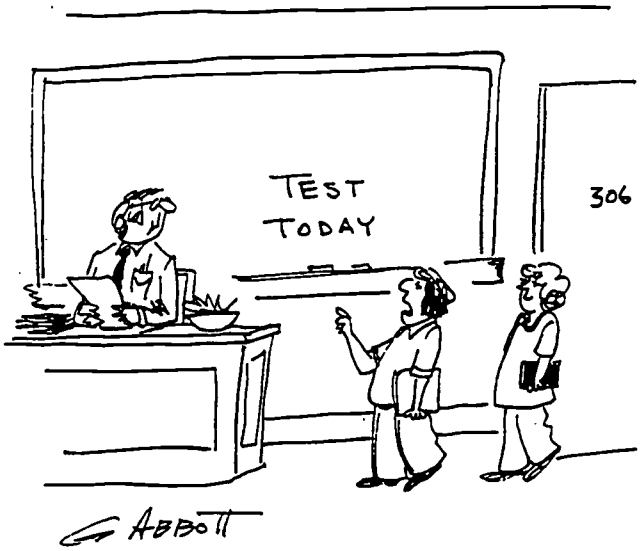
Tests given in class and tests and other exercises assigned for homework are also important means of promoting feedback. A good test can be an occasion for learning. It is better to have frequent short tests than infrequent long ones. Any new learning should first be tested within about a week of a first encounter, but more frequent tests are counterproductive. The quality of the test items — that is, their relevance to the main learning aims and their clear communication to the pupil — requires scrutiny as well. Good questions are hard to generate, and teachers should collaborate and draw on outside sources to collect such questions.

Given questions of good quality, it is essential to ensure the quality of the feedback. Research studies have shown that, if pupils are given only marks or grades, they do not benefit from the feedback. The worst scenario is one in which some pupils who get low marks this time also got low marks last time and come to expect to get low marks next time. This cycle of repeated failure becomes part of a shared belief between such students and their teacher. Feedback has been shown to improve learning when it gives each pupil specific guidance on strengths and weaknesses, preferably without any overall marks. Thus the way in which test results are reported to pupils so that they can identify their own strengths and weaknesses is critical. Pupils must be given the means and opportunities to work with evidence of their difficulties. For formative purposes, a test at the end of a unit or teaching module is pointless; it is too late to work with the results. We conclude that *the feedback on tests, seatwork, and homework should give each pupil guidance on how to improve, and each pupil must be given help and an opportunity to work on the improvement.*

All these points make clear that there is no one simple way to improve formative assessment. What is common to them is that a teacher's approach should start by being realistic and confronting the question "Do I really know enough about the understanding of my pupils to be able to help each of them?"

Much of the work teachers must do to make good use of formative assessment can give rise to difficulties. Some pupils will resist attempts to change accustomed routines, for any such change is uncomfortable, and emphasis on the challenge to think for yourself (and not just to work harder) can be threatening to many. Pupils cannot be expected to believe in the value of changes for their learning before they have experienced the benefits of such changes. Moreover, many of the initiatives that are needed take more class time, particularly when a central purpose is to change the outlook on learning and the working methods of pupils. Thus teachers have to take risks in the belief that such investment of time will yield rewards in the future, while "delivery" and "coverage" with poor understanding are pointless and can even be harmful.

Teachers must deal with two basic issues that are the source of many of the problems associated with changing to a system of formative assessment. The first is *the nature of each teacher's beliefs about learning.* If the teacher assumes that knowledge is to be transmitted and learned, t understanding will develop later, and that clarity of exposition accompanied by rewards for patient reception are the essentials of good teaching, then formative assessment is hardly necessary. However, most teachers accept the wealth of evi-



"It has been said that a fool can ask more questions than a wise man can answer."

dence that this transmission model does not work, even when judged by its own criteria, and so are willing to make a commitment to teaching through interaction. Formative assessment is an essential component of such instruction. We do not mean to imply that individualized, one-on-one teaching is the only solution; rather we mean that what is needed is a classroom culture of questioning and deep thinking, in which pupils learn from shared discussions with teachers and peers. What emerges very clearly here is the indivisibility of instruction and formative assessment practices.

The other issue that can create problems for teachers who wish to adopt an interactive model of teaching and learning relates to *the beliefs teachers hold about the potential of all their pupils for learning*. To sharpen the contrast by overstating it, there is on the one hand the "fixed I.Q." view — a belief that each pupil has a fixed, inherited intelligence that cannot be altered much by schooling. On the other hand, there is the "untapped potential" view — a belief that starts from the assumption that so-called ability is a complex of skills that can be learned. Here, we argue for the underlying belief that all pupils can learn more effectively if one can clear away, by sensitive handling, the obstacles to learning, be they cognitive failures never diagnosed or damage to personal confidence or a combination of the two. Clearly the truth lies between these two extremes, but the evidence is that *ways of managing formative assessment that work with the assumptions of "untapped potential" do help all pupils to learn and can give particular help to those who have previously struggled.*

Policy and Practice

Changing the policy perspective. The assumptions that drive national and state policies for assessment have to be called into question. The promotion of testing as an important component for establishing a competitive market in education can be very harmful. The more recent shifting of emphasis toward setting targets for all, with assessment providing a touchstone to help check pupils' attainments, is a more mature position. However, we would argue that *there is a need now to move further, to focus on the inside of the "black box" and so to explore the potential of assessment to raise standards directly as an integral part of each pupil's learning work.*

It follows from this view that several changes are needed. First, policy ought to start with a recognition that the prime locus for raising standards is the classroom, so that the overarching priority has to be the promotion and support of change within the classroom. Attempts to raise standards by reforming the inputs to and measuring the outputs from the black box of the classroom can be helpful, but they are not adequate on their own. Indeed, their helpfulness can be judged only in light of their effects in classrooms.

The evidence we have presented here establishes that a clearly productive way to start implementing a classroom-focused policy would be to improve formative assessment. This same evidence also establishes that in doing so we would not be concentrating on some minor aspect of the business of teaching and learning. Rather, we would be concentrating on several essential elements: the quality of teacher/pupil interactions, the stimulus and help for pupils to take active responsibility for their own learning, the particular help needed to move pupils out of the trap of "low achievement," and the development of the habits necessary for all students to become lifelong learners. Improvements in formative assessment, which are within the reach of all teachers, can contribute substantially to raising standards in all these ways.

Four steps to implementation. If we accept the argument outlined above, what needs to be done? The proposals outlined below do not follow directly from our analysis of assessment research. They are consistent with its main findings. but they also call on more general sources for guidance.[16]

At one extreme, one might call for more research to find out how best to carry out such work; at the other, one might call for an immediate and large-scale program, with new guidelines that all teachers should put into practice. Neither of these alternatives is sensible: while the first is unnecessary because enough is known from the results of research, the second would be unjustified because not enough is known about classroom practicalities in the context of any one country's schools.

Thus the improvement of formative assessment cannot be a simple matter. There is no quick fix that can alter existing practice by promising rapid rewards. On the contrary, if the substantial rewards promised by the research evidence are to be secured, each teacher must find his or her own ways of incorporating the lessons and ideas set out above into his or her own patterns of classroom work and into the cultural norms and expectations of a particular school community.[17] This process is a relatively slow one and takes place through sustained programs of professional development and support. This fact does not weaken the message here; indeed, it should be seen as a sign of its authenticity, for lasting and fundamental improvements in teaching and learning must take place in this way. A recent international study of innovation and change in education, encompassing 23 projects in 13 member countries of the Organisation for Economic Co-operation and Development, has arrived at exactly the same conclusion with regard to effective policies for change.[18] Such arguments lead us to propose a four-point scheme for teacher development.

1. *Learning from development.* Teachers will not take up ideas that sound attractive, no matter how extensive the research base, if the ideas are presented as general principles that leave the task of translating them into everyday practice entirely up to the teachers. Their classroom lives are too busy and too fragile for all but an outstanding few to undertake such work. What teachers need is a variety of living examples of implementation, as practiced by teachers with whom they can identify and from whom they can derive the confidence that they can do better. They need to see examples of what doing better means in practice.

So changing teachers' practice cannot begin with an extensive program of training for all; that could be justified only if it could be claimed that we have enough "trainers" who know what to do, which is certainly not the case. The essential first step is to set up a small number of local groups of schools — some primary, some secondary, some inner-city. some from outer suburbs, some rural — with each school committed both to a school-based development of formative assessment and to collaboration with other schools in its local group. In such a process, the teachers in their classrooms will be working out the answers to many of the practica questions that the evidence presented here cannot answer. They will be reformulating the issues, perhaps in relation to fundamental insights and certainly in terms that make sense to their peers in other class-

rooms. It is also essential to carry out such development in a range of subject areas, for the research in mathematics education is significantly different from that in language, which is different again from that in the creative arts.

The schools involved would need extra support in order to give their teachers time to plan the initiative in light of existing evidence, to reflect on their experience as it develops, and to offer advice about training others in the future. In addition, there would be a need for external evaluators to help the teachers with their development work and to collect evidence of its effectiveness. Video studies of classroom work would be essential for disseminating findings to others.

2. *Dissemination.* This dimension of the implementation would be in low gear at the outset — offering schools no more than general encouragement and explanation of some of the relevant evidence that they might consider in light of their existing practices. Dissemination efforts would become more active as results and resources became available from the development program. Then strategies for wider dissemination — for example, earmarking funds for inservice training programs — would have to be pursued.

We must emphasize that this process will inevitably be a slow one. To repeat what we said above, *if the substantial rewards promised by the evidence are to be secured, each teacher must find his or her own ways of incorporating the lessons and ideas that are set out above into his or her own patterns of classroom work.* Even with optimum training and support, such a process will take time.

3. *Reducing obstacles.* All features in the education system that actually obstruct the development of effective formative assessment should be examined to see how their negative effects can be reduced. Consider the conclusions from a study of teachers of English in U.S. secondary schools.

> Most of the teachers in this study were caught in conflicts among belief systems and institutional structures, agendas, and values. The point of friction among these conflicts was assessment, which was associated with very powerful feelings of being overwhelmed, and of insecurity, guilt, frustration, and anger. . . . This study suggests that assessment, as it occurs in schools, is far from a merely technical problem. Rather, it is deeply social and personal.[19]

The chief negative influence here is that of short external tests. Such tests can dominate teachers' work, and, insofar as they encourage drilling to produce right answers to short, out-of-context questions, they can lead teachers to act against their own better judgment about the best ways to develop the learning of their pupils. This is not to argue that all such tests are unhelpful. Indeed, they have an important role to play in securing public confidence in the accountability of schools. For the immediate future, what is needed in any development program for formative assessment is to study the interactions between these external tests and formative assessments to see how the models of assessment that external tests can provide could be made more helpful.

All teachers have to undertake some summative assessment. They must report to parents and produce end-of-year reports as classes are due to move on to new teachers. However, the task of assessing

pupils summatively for external purposes is clearly different from the task of assessing ongoing work to monitor and improve progress. Some argue that these two roles are so different that they should be kept apart. We do not see how this can be done, given that teachers must have some share of responsibility for the former and must take the leading responsibility for the latter.[20] However, teachers clearly face difficult problems in reconciling their formative and summative roles, and confusion in teachers' minds between these roles can impede the improvement of practice.

The arguments here could be taken much further to make the case that teachers should play a far greater role in contributing to summative assessments for accountability. One strong reason for giving teachers a greater role is that they have access to the performance of their pupils in a variety of contexts and over extended periods of time.

This is an important advantage because sampling pupils' achievement by means of short exercises taken under the conditions of formal testing is fraught with dangers. It is now clear that performance in any task varies with the context in which it is presented. Thus some pupils who seem incompetent in tackling a problem under test conditions can look quite different in the more realistic conditions of an everyday encounter with an equivalent problem. Indeed, the conditions under which formal tests are taken threaten validity because they are quite unlike those of everyday performance. An outstanding example here is that collaborative work is very important in everyday life but is forbidden by current norms of formal testing.[21] These points open up wider arguments about assessment systems as a whole — arguments that are beyond the scope of this article.

4. *Research.* It is not difficult to set out a list of questions that would justify further research in this area. Although there are many and varied reports of successful innovations, they generally fail to give clear accounts of one or another of the important details. For example, they are often silent about the actual classroom methods used, the motivation and experience of the teachers, the nature of the tests used as measures of success, or the outlooks and expectations of the pupils involved.

However, while there is ample justification for proceeding with carefully formulated projects, we do not suggest that everyone else should wait for their conclusions. Enough is known to provide a basis for active development work, and some of the most important questions can be answered only through a program of practical implementation.

Directions for future research could include a study of the ways in which teachers understand and deal with the relationship between their formative and summative roles or a comparative study of the predictive validity of teachers' summative assessments versus external test results. Many more questions could be formulated, and it is important for future development that some of these problems be tackled by basic research. At the same time, experienced researchers would also have a vital role to play in the evaluation of the development programs we have proposed.

Are We Serious About Raising Standards?

The findings summarized above and the program we have outlined have implications for a variety of responsible agencies. However, it is the responsibility of governments to take the lead. It would be premature and out of order for us to try to consider the relative roles in such an effort, although success would clearly depend on cooperation among government agencies, academic researchers, and school-based educators.

The main plank of our argument is that standards can be raised only by changes that are put into direct effect by teachers and pupils in classrooms. There is a body of firm evidence that formative assessment is an essential component of classroom work and that its development can raise standards of achievement. We know of no other way of raising standards for which such a strong prima facie case can be made. Our plea is that national and state policy makers will grasp this opportunity and take the lead in this direction.

1. James W. Stigler and James Hiebert, "Understanding and Improving Classroom Mathematics Instruction: An Overview of the TIMSS Video Study," *Phi Delta Kappan*, September 1997, pp. 19-20.

2. There is no internationally agreed-upon term here. "Classroom evaluation," "classroom assessment," "internal assessment," "instructional assessment," and "student assessment" have been used by different authors, and some of these terms have different meanings in different texts.

3. Paul Black and Dylan Wiliam, "Assessment and Classroom Learning," *Assessment in Education*, March 1998, pp. 7-74.

4. Lynn S. Fuchs and Douglas Fuchs, "Effects of Systematic Formative Evaluation: A Meta-Analysis," *Exceptional Children*, vol. 53, 1986, pp. 199-208.

5. See Albert E. Beaton et al., *Mathematics Achievement in the Middle School Years* (Boston: Boston College, 1996).

6. Lynn S. Fuchs et al., "Effects of Task-Focused Goals on Low-Achieving Students with and Without Learning Disabilities," *American Educational Research Journal*, vol. 34, 1997, pp. 513-43.

7. OFSTED (Office for Standards in Education), *Subjects and Standards: Issues for School Development Arising from OFSTED Inspection Findings 1994-5: Key Stages 3 and 4 and Post-16* (London: Her Majesty's Stationery Office, 1996), p. 40.

8. Nicholas Daws and Birendra Singh, "Formative Assessment: To What Extent Is Its Potential to Enhance Pupils' Science Being Realized?," *School Science Review*, vol. 77, 1996, p. 99.

9. Clement Dassa, Jesús Vazquez-Abad, and Djavid Ajar, "Formative Assessment in a Classroom Setting: From Practice to Computer Innovations," *Alberta Journal of Educational Research*, vol. 39, 1993, p. 116.

10. D. Monty Neill, "Transforming Student Assessment," *Phi Delta Kappan*, September 1997, pp. 35-36.

11. *Task Group on Assessment and Testing: A Report* (London: Department of Education and Science and the Welsh Office, 1988).

12. Richard Daugherty, *National Curriculum Assessment: A Review of Policy, 1987-1994* (London: Falmer Press, 1995).

13. Terry A. Russell, Anne Qualter, and Linda McGuigan, "Reflections on the Implementation of National Curriculum Science Policy for the 5-14 Age Range: Findings and Interpretations from a National Evaluation Study in England," *International Journal of Science Education*, vol. 17, 1995, pp. 481-92.

14. Phillipe Perrenoud, "Towards a Pragmatic Approach to Formative Evaluation," in Penelope Weston, ed., *Assessment of Pupils' Achievement: Motivation and School Success* (Amsterdam: Swets and Zeitlinger, 1991), p. 92.

15. D. Royce Sadler, "Formative Assessment and the Design of Instructional Systems," *Instructional Science*, vol. 18, 1989, pp. 119-44.

16. Paul J. Black and J. Myron Atkin, *Changing the Subject: Innovations in Science, Mathematics, and Technology Education* (London: Routledge for the Organisation for Economic Co-operation and Development, 1996); and Michael G. Fullan, with Suzanne Stiegelbauer, *The New Meaning of Educational Change* (London: Cassell, 1991).

17. See Stigler and Hiebert, pp. 19-20.

18. Black and Atkin, op. cit.

19. Peter Johnston et al., "Assessment of Teaching and Learning in Literature-Based Classrooms," *Teaching and Teacher Education*, vol. 11, 1995, p. 359.

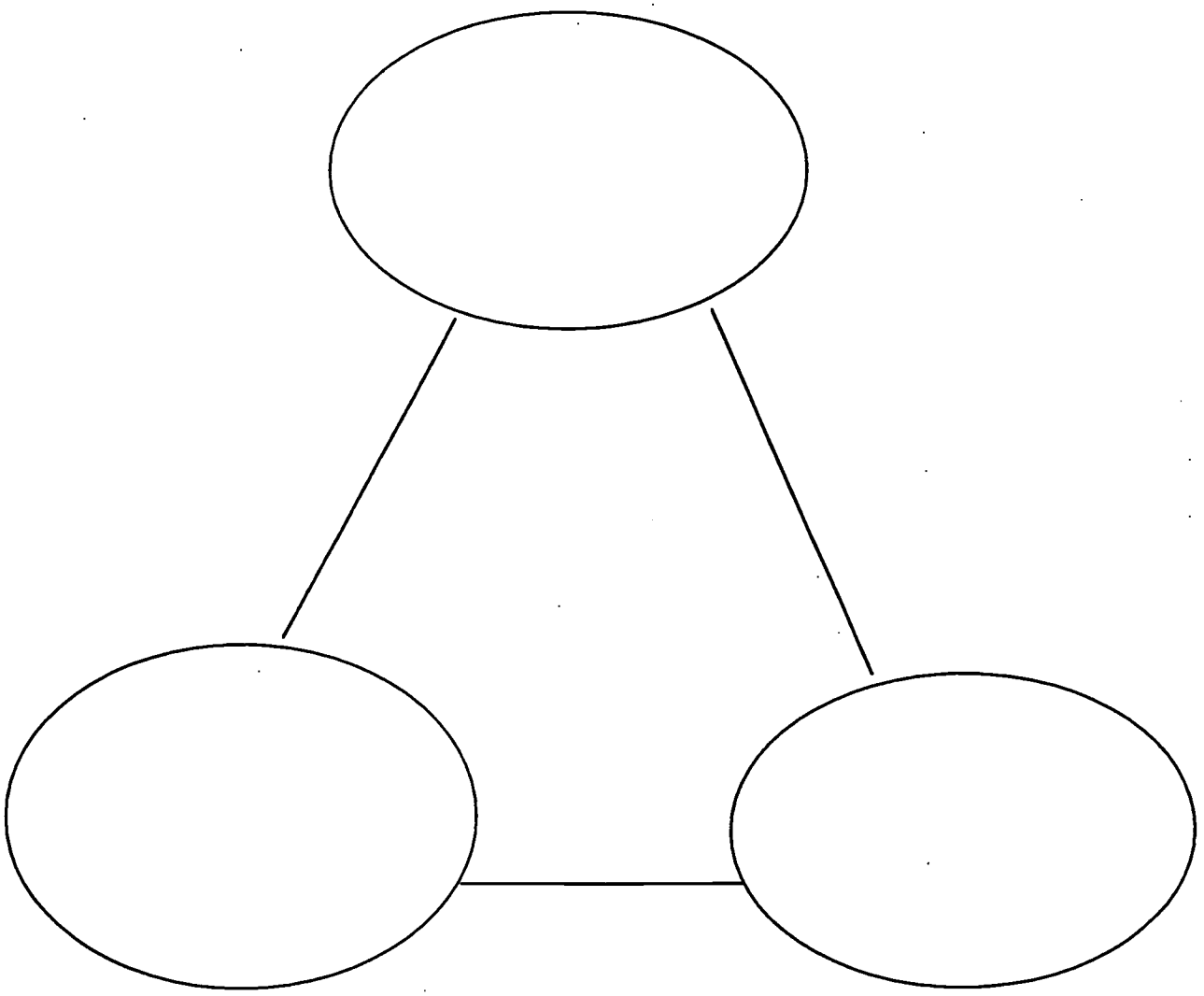
20. Dylan Wiliam and Paul Black, "Meanings and Consequences: A Basis for Distinguishing Formative and Summative Functions of Assessment," *British Educational Research Journal*, vol. 22, 1996, pp. 537-48.

21. These points are developed in some detail in Sam Wineburg, "T. S. Eliot, Collaboration, and the Quandaries of Assessment in a Rapidly Changing World," *Phi Delta Kappan*, September 1997, pp. 59-65.

Checking for understanding – Part I

Directions:

Label the parts of the assessment cycle. Can you do this without looking back in the manual? Check page 19 in the manual and make any additions or corrections to your model. In your own mind, can you give an example for each part of the model?



Checking for understanding – Part II

Define the terms used to relate assessment as a process.

TERM	YOUR OWN DEFINITION	HOW DOES THIS TERM RELATE TO THE ASSESSMENT CYCLE FOUND ON PAGE 19 IN CHAPTER ONE'S TEXT?
Example: **Formative Assessment**	*Formative assessment reflects what I am doing with my students as part of the instructional process. It is ongoing and may or may not reflect a student grade. Formative assessment helps me give students quick feedback for improvement and guides my next steps in instruction.*	*Formative assessment fits into specific areas of this cycle as related to the purposes of assessment, the information needed, the feedback given, the inferences made , and communication with students. Formative assessment constantly reminds my students and me of what I really want my students to know and be able to do.*
Summative Assessment		
Learning Targets		
Diagnosis (related to assessment)		
Feedback		
Inferences		

Clarifying Learning Targets

Chapter Two

Goals:

The focus of this section is to develop an understanding of the importance of clarity of what students must know and be able to do. As you finish this section you will:

1. Clarify what learning targets are.
2. Conceptualize levels of learning targets on a continuum from more comprehensive to most specific.
3. Clarify different categories of learning targets (knowledge, reasoning, skills, products, attitudes).
4. Clarify and determine appropriate instructional activities that will best teach students the targets that will be assessed.

Estimated time for this module: 3 hours

Assigned readings:

Chapter 2 in the manual, "Clarifying Learning Targets," pages 35--42

Questions for discussion and reflection:

1. How would you explain to a new teacher the difference between using the *Standard Course of Study* as your guide to determine what you teach versus using the textbook pages to determine what you teach?
2. What does the following statement mean? "When you are planning for instruction you should plan for assessment at the same time."
3. What is the difference in assessing an activity and assessing student understanding of a learning target?

Required activities:

- Activity 1: Clarifying Learning Targets
- Activity 2: Creating the Big Picture

Optional activities (Choose one):

- Activity 3: Story Swap
- Activity 4: Discussion of Activity 1

Checking for understanding:
Draft a letter to parents

Reflecting on what was learned:

1. How does developing a conceptual road map or knowing very clearly what I'm responsible for teaching make me a more effective teacher?

2. What are the relationships among my *Standard Course of Study*, my pacing guide, my textbook, and other resources?

Complete this checklist when you have finished this module:

_____ Read pp. 35-42 in the manual

_____ Discussed "Questions for discussion and reflection"

_____ Completed "Checking for Understanding"

_____ Completed Activity 1 – Clarifying Learning Targets

_____ Completed Activity 2 – Creating the Big Picture

_____ Completed one activity of your choice

_____ Reflected on what was learned

Activity 1

CLARIFYING LEARNING TARGETS

Directions: Select one topic your class is studying or will soon study, and, using the *Standard Course of Study*, determine the goals and objectives that need to be learned: Complete the following chart, thinking about using learning targets as your guide for instruction.

Topic/unit/subject for this study: _____

<table>
<tr><th>Guiding Question/s</th><th colspan="5">Your Comment/s</th></tr>
<tr><td>1. Write down the basic goals that you want your students to learn in this unit. Use the Standard Course of Study to help you identify those goals.

(These goals are also called the learning targets)</td><td colspan="5"></td></tr>
<tr><td>2. Are there additional learning targets that are necessary for mastering the topic? For example: working in groups or technical writing objectives</td><td colspan="5"></td></tr>
<tr><td>3. What are the important objectives that are smaller components or pieces of the broad goals? These should read: The student will be able to:</td><td colspan="5"></td></tr>
<tr><td>4. Look at your goals and objectives and note how many of them fall into each category. (If you need to, review these categories on page 39 of the text.)</td><td>Knowledge and Information</td><td>Reasoning and problem solving</td><td>Skills and processes</td><td>Products and applications</td><td>Attitudes and dispositions</td></tr>
</table>

Continued on the next page

5. Which learning targets are the most important for this topic? Prioritize the goals (learning targets).	
6. Describe what students should know and be able to do to provide evidence that they have attained these goals.	
7. What connection to previous learning, the real world, or other topics would be useful?	
8. How will you explain, model, illustrate, or describe these targets so that they are clear to you and to your students? (Include criteria for quality performance.)	
9. What are the likely misconceptions your students may have with the learning targets for this topic?	
10. How much time do you think your students will need to attain these learning targets?	

(Keep these activity sheets for reference in Chapter 3)

Activity 2

CREATING THE BIG PICTURE

Directions:
Create a visual representation of a course you teach and reflect the major things you want your students to know and be able to do.

This visual may be something you give to parents as a guide to your course of study or a large poster that is placed in your room to show the students what they will learn.

Activity 3

STORY SWAP

Directions:

Think of a story to share with colleagues about an experience in which the directions or ultimate goals were unclear to you or a student.

For example: Following the elections when the student council met for the first time, the sponsors gave the new officers job descriptions. One advisor asked the secretary, whose job description included handling correspondence and "taking minutes" in the meetings, to get a list of all council members with home telephone numbers. After the meeting, the newly elected secretary brought the list to the advisor and said, "Thanks for giving me the job description. I wasn't sure what the secretary was supposed to do." When the advisor looked at the list of names and phone numbers, at the bottom of the page he saw written: "*Minutes 7:05 – 7:45.*"

Activity 4

DISCUSSION OF ACTIVITY 1

Directions:

Share with colleagues teaching the same grade or a similar course your answers to the guiding questions for activity 1, "Clarifying Learning Targets." Does everyone agree on what the goals and objectives are and how students demonstrate that they have attained those goals and objectives? What are some of the insights you gained from activity 1? Did you gain other insights from your colleagues' ideas?

CHECKING FOR UNDERSTANDING

Directions:
Draft a letter to parents that would describe what a student is expected to know for one subject or course. Include a description of what quality performance looks like. You may wish to include the graphic you created in Activity 2.

156

Using Multiple Assessment Strategies

Chapter Three

Goals:

The focus of this section is to explore various assessment strategies. As you finish this section you will:

1. Know the five major assessment strategies (categories) for assessing student learning and know the various assessment methods associated with each strategy.
2. Match the kinds of learning targets with the assessment methods that will give the best evidence or information of what students know and are able to do.
3. Understand the importance of using assessment variety for better information about student thinking and reasoning.

Estimated time for this module: 3 hours

Assigned readings:

Chapter 3 in the manual, "Using Multiple Assessment Strategies," pages 43 - 62. (You will need a copy of your *Standard Course of Study* for this section.)

Questions for discussion and reflection:

1. What assessment methods do you use with your students most often? Why?
2. What types of information do you gain from the different assessment strategies? What are the advantages and disadvantages of each strategy?
3. How does it help the student when you use a variety of assessment methods?

Required activities:

- Activity 1 – Using Multiple Assessment Methods
- Activity 2 – Analogy: Learning as a Road Map
- Activity 3 – Analyzing Assessment Methods

Checking for understanding:

Open-ended questions

Reflecting on what was learned:

1. How may using multiple assessment methods result in improved student learning and achievement?

2. What is one assessment strategy that I have not used frequently and that I might consider implementing?

Complete this checklist when you have finished this module:

_____ Read pp. 43 - 62 in the manual

_____ Discussed "Questions for discussion and reflection"

_____ Completed Activity 1 – Using Multiple Assessment Methods

_____ Completed Activity 2 – Analogy: Learning as a Road Map

_____ Completed Activity 3 – Analyzing Assessment Methods

_____ Completed "Checking for Understanding"

_____ Reflected on what was learned

Activity 1

Using Multiple Assessment Methods

Directions: Refer to Activity 1 in Chapter 2 and use the information you collected and recorded. You will use the same goals and objectives (pick one or two of the major ones) to take a closer look at the appropriate assessment methods and strategies you will consider for assessing these learning targets.

Think about	Goal 1	Goal 2
What goals have you chosen?		
What categories of learning targets are you assessing?		
What assessment methods may be best to use and why? (Refer to the five basic strategies and define the actual method you will use.)		
If you use this method, what will it tell you about your students?		
What problems might your students have with this assessment method?		
What preparation would help students demonstrate their learning through this method of assessment?		

Activity 2:

Analogy: Learning as a Roadmap

Directions:

Read this story and then discuss the questions at the end of the story.

A married couple, Sarah and Frank, both of whom are teachers, love to travel in the summer. They prefer to take road trips, visiting interesting places they have never seen. Sarah is always the designated driver and Frank always reads the map and gives directions. After a number of these trips, Sarah, who is a third grade teacher, tells Frank:

"Even though I'm always the driver, I couldn't possibly retrace any of the trips we have taken. I just follow your directions, turning when you tell me to.

"That's like my students. Even though they're doing the work, they rely on me for directions, following orders but not understanding where they are or where they are going. I think they need to become map readers!"

As teachers, we have always used a variety of assessment methods. Where some teachers may fall short is in helping students know what learning targets are and what criteria determine success in these assessments.

Discuss these questions on how this story relates to the classroom:

1. How is the destination of Sarah and Frank's trip like a learning target?
2. Can students reach a destination without understanding how they arrived there? Explain.

Activity 3

Analyzing Assessment Methods

Directions: Take the last four assessments you gave your students and analyze them according to assessment method and category of learning target.

ASSESSMENT METHODS USED	CATEGORIES OF LEARNING TARGETS	DID THE METHOD MATCH THE INTENDED TARGET? HOW DO YOU KNOW?
1.		
2.		
3.		
4.		

Did you use a variety of assessment methods with the four assessments you used with your students? Why is this important?

Checking for Understanding

1. What are five assessment strategies? Give two methods that exemplify each strategy?

2. When would a teacher use selected response?

3. What are some criteria for effective performance assessments or tasks?

4. When are observations helpful?

5. When would you use a checklist?

Making Decisions and Taking Action

Chapter 4

Goals:

The focus of this section is to use assessment data to make instructional decisions and give feedback to students. As you finish this section you will:

1. Describe the importance of quality evidence and appropriate methods in making judgments about student achievement.
2. Describe the use of assessment information in making decisions and taking action to help students achieve
3. Know the many different kinds of decisions that assessment data can inform.
4. Discuss how the quality of decisions you make is influenced by the quality of data you have.

Estimated time for this module: 5 hours

Assigned readings:

Chapter 4 in the manual, "Making Decisions and Taking Action," pages 63 - 76

Questions for discussion and reflection

1. After you have completed an assessment with your students, how do you use the information about individual students to make decisions about the class as a whole?
2. What other types of decisions do you make using assessment information? Explain.
3. How do you provide feedback to students and what is the nature of that feedback about individual student learning?
4. Having studied the assessment and instructional cycle and reflected on learning targets and assessment strategies, what decisions might you make in your classroom to make assessment more a part of your ongoing daily classroom routine?

Required activities:

- Activity 1- The Camping Trip
- Activity 2- Looking at your Students' Work
- Activity 3 - Critique of Assessment Environment. *If your students are old enough to read this form, allow them to complete the same form and then check to see how closely your perceptions match theirs.*
- Activity 4 – Video Segment Review

Checking for understanding:

Short response questions

Reflecting on what was learned:

1. How confident am I that the assessments I am using in the classroom really measure what I want to measure? Explain.

2. How do the instructional decisions that I make about individual students or groups influence their achievement at the end of the year?

Complete this checklist when you have finished this module.

_____ Read pp. 63-76 in manual

_____ Discussed "Questions for discussion and reflection"

_____ Completed Activity 1 -The Camping Trip

_____ Completed Activity 2 - Looking at your Students' Work

_____ Completed Activity 3 - Critique of Assessment Environment

_____ Completed Activity 4 – Video Segment Review

_____ Completed "Checking for Understanding"

_____ Reflected on what was learned

Activity 1

The Camping Trip

(This activity is adapted from Improving Classroom Assessment: A Toolkit for Professional Developers, Toolkit 98, Appendix B, Sample b.4, SERVE.)

This activity involves judging the quality of fifth graders' work on an open-ended mathematics prompt. Cut apart the different student work samples and sort them into three stacks (*strong*, *medium*, and *weak* responses). Then, in pairs, discuss the sorting and reach a consensus on which papers belong in which stacks. Fill in the chart below to describe the features that make stack 1 different from stack 2 and stack 2 different from stack 3. (In this way, you are beginning to delineate criteria for quality responses to open-ended mathematics questions.)

	High	Medium	Low
Record which samples are in each category.			
What are the criteria for placement in this category?			

Discuss your categorizations and the criteria that underlie them. Did you comment on accuracy, clarity of explanation, logical reasoning, and appropriateness of mathematics for the given prompt?

Sample 1

14. A group of 8 people are all going camping for three days and need to carry their own water. They read in a guide book that 12.5 liters are needed for a party of 5 people for 1 day. Based on the guide book, what is the minimum amount of water the 8 people should carry all together?

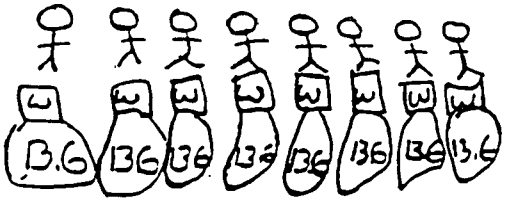
Explain your answer.

If 12.5 liters are needed for 5 people for 1 day well then 45 liters will be needed for three days for 8 people because for 1 person it was 1.5 which all together equaled 12.5 for 1 day, for 8 people to go on a 1 day trip would have ~~to~~ bring 45 liters.

Sample 2

14. A group of 8 people are all going camping for three days and need to carry their own water. They read in a guide book that 12.5 liters are needed for a party of 5 people for 1 day. Based on the guide book, what is the minimum amount of water the 8 people should carry all together?

Explain your answer.



$$\begin{array}{r} 12.5 \\ 3 \\ +\ 8 \\ \hline 13.2 \end{array}$$

liters of water
13.6 for 8 people to bring
camping for 3 days

Sample 3

14. A group of 8 people are all going camping for three days and need to carry their own water. They read in a guide book that 12.5 liters are needed for a party of 5 people for 1 day. Based on the guide book, what is the minimum amount of water the 8 people should carry all together?

Explain your answer. 60 Liters

I came up with that by getting how much 1 person needed Then how much 8 people needed that was 20 and then I multaplied that by three for three days and came up with sixty

Sample 4

14. A group of 8 people are all going camping for three days and need to carry their own water. They read in a guide book that 12.5 liters are needed for a party of 5 people for 1 day. Based on the guide book, what is the minimum amount of water the 8 people should carry all together?

Explain your answer. 20. liters for 8 people one day

2.5 for each person

12.5
÷ 5
2.5

12.5
+ 2.5
15.0
+ 2.5
17.5
+ 2.5
20.0

Sample 5

14. A group of 8 people are all going camping for three days and need to carry their own water. They read in a guide book that 12.5 liters are needed for a party of 5 people for 1 day. Based on the guide book, what is the minimum amount of water the 8 people should carry all together?

Explain your answer.

$5 \overline{)12.5}$ liters / people → 2.5 liters/person

2.5 l/person × 8 people = 20 liters/day

20 liters/day × 3 days = 60 liters in all

I divided 12.5 liters ÷ 5 people = 2.5 liters/person. I did that so that I could take 2.5 liters × 8 people = 20 liters/day. Now I need to multiply 20 liters/day x 3 days = 60 liters to last the whole camping trip. 60 Liters in all.

Sample 6

12.5 litters × 2 = 25.0 litters

25.0 litters × 8 people = 200.0 litters

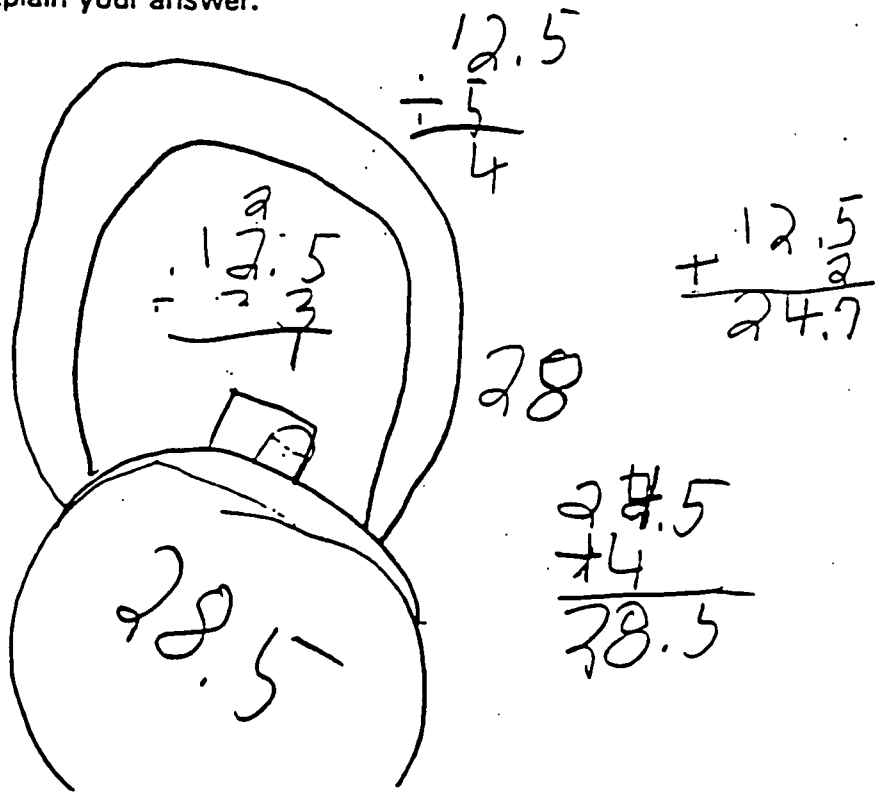
200.0 litters × 3 days = 600.0 litters are needed for 8 people for 3 days

168

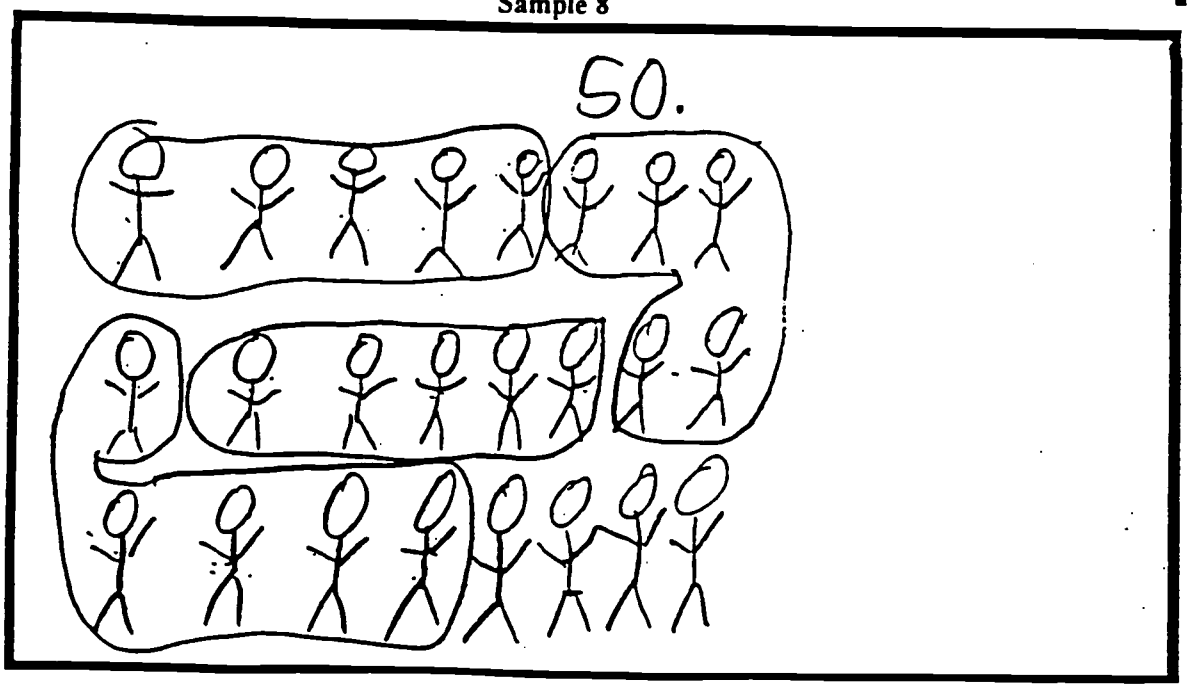
Sample 7

14. A group of 8 people are all going camping for three days and need to carry their own water. They read in a guide book that 12.5 liters are needed for a party of 5 people for 1 day. Based on the guide book, what is the minimum amount of water the 8 people should carry all together?

Explain your answer.



Sample 8



Sample 9

14. A group of 8 people are all going camping for three days and need to carry their own water. They read in a guide book that 12.5 liters are needed for a party of 5 people for 1 day. Based on the guide book, what is the minimum amount of water the 8 people should carry all together?

Explain your answer.

2.5 liters of water apiece
5 people | 12.5 lieters of water for a group of 5 people.

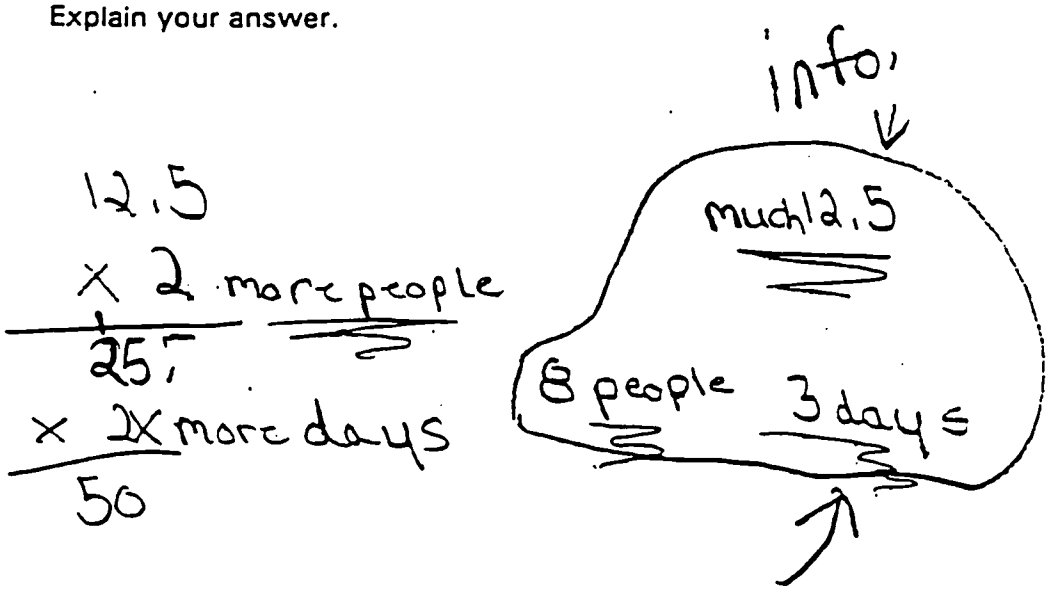
2.5 liters each of water
x 8 people
20 lieters of water needs to be token at 2.5 lieters each

170

Sample 10

14. A group of 8 people are all going camping for three days and need to carry their own water. They read in a guide book that 12.5 liters are needed for a party of 5 people for 1 day. Based on the guide book, what is the minimum amount of water the 8 people should carry all together?

Explain your answer.



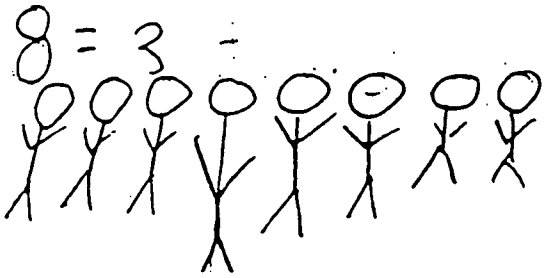
12.5
× 2 more people
25
× 2X more days
50

First e gathered some important info. 5 peopl for one day was 12.5 e add two people by timing to equel 25. then e times it by two again for the day e new that one day was all ready carried 50 liters

Sample 11

14. A group of 8 people are all going camping for three days and need to carry their own water. They read in a guide book that 12.5 liters are needed for a party of 5 people for 1 day. Based on the guide book, what is the minimum amount of water the 8 people should carry all together?

Explain your answer.



12.5 ÷ 5 = 2.5

2.5 × 4 = 10

10 + 50 = 60

Sample 12

14. A group of 8 people are all going camping for three days and need to carry their own water. They read in a guide book that 12.5 liters are needed for a party of 5 people for 1 day. Based on the guide book, what is the minimum amount of water the 8 people should carry all together?

Explain your answer. They can bring enough water for 5 people for 6 days which 75 liters which should be enough

Activity 2

Looking at your Students' Work

This activity involves judging the *quality* of the work of students in your classes. For this activity, work in groups of no more than three.

Directions: Each teacher should bring a sample set of papers from a recent assignment, including the directions (prompt) to the class. This set should be 8 samples that include a full range of performance levels. The papers should not have scores or student names on them. The process will be facilitated if the papers are copied so that each person has a set. For each set of papers, sort the work into stacks of high, medium, and low performance and discuss how the work in each stack is different.

How would a score be obtained for these papers? How would the scores relate to the criteria you established in sorting the papers?

Are there patterns in the types of errors individual students are making? Could these errors have been avoided by clearer instructions.

Are there patterns in the types of errors the students as a class are making?

Can you see what students understand clearly and can perform proficiently?

Can you spot gaps in their learning or misconceptions?

Activity 3

Critique of the Assessment Environment in this Class

Directions:
Read each question and check the appropriate box. This data collection sheet is appropriate both for teachers and middle and high school students.

Assessments in this classroom:	Check the appropriate box		
	Never	Sometimes	Most of the time
1. Assess students on clear goals			
2. Ask students to go beyond simple recall of facts			
3. Give clear criteria for quality performance to students prior to the assessment			
4. Allow students an opportunity to self assess or revise work			
5. Allow students an opportunity to peer assess and help each other improve			
6. Reveal levels of proficiency (not just correct or incorrect responses)			
7. Are used by the teacher to improve and adjust instructional strategies			
8. Provide clear and concise instructions to students			
9. Are interesting to do			
10. Reflect what the state, district or school value in student learning			
11. Are designed as free as possible of cultural, ethnic, or gender stereotypes			
12. Allow adequate time for completion			
13. Are worth doing			
14. Have an appropriate level of difficulty			
15. Allow students to know how a score will be determined			

What actions should be taken to improve assessments in this classroom?

Activity 4

Video Segment Review

Directions:
Watch the Annenberg videotape of the grade span (K-4, 5-8, and 9-12) which you teach. There are introductory materials in the manual that will acquaint you with the videotape before you view the segment appropriate to your grade span.

Think about and discuss: How are teachers in this videotape using assessment data to make decisions?

Checking for Understanding

Give at least three examples that illustrate:
The difference between high and low inferences:
The issues and ideas you consider when interpreting assessment data:
The criteria for quality evidence:
The kinds of decisions teachers make using assessment data:

Documenting and Communicating

Chapter 5

Goals:

The focus of this chapter is why documentation and communication of assessment information helps students take greater responsibility for their learning. As you finish this chapter you will:

1. Understand how efficient ways to document and manage assessment information are important.
2. Understand how to provide feedback for learning and teaching.
3. Discuss the importance of clear and effective assessment information for all targeted audiences.
4. Consider communication options for clarifying student expectations that encourage students to take greater responsibility for their own learning.
5. Understand what information is needed to determine a proficiency level for the intended learning.

Estimated time for this module: 3 hours

Special Note: This chapter contains a great deal of information that covers four very important topics: portfolio assessment, grading and reporting issues, the use of rubrics, and the use of student-led parent conferences. The context of this chapter is not meant to address these topics in depth but only to give basic information for gaining a fundamental literacy level for each topic.

Assigned readings:

Chapter 5 in the manual, "Documenting and Communicating," pages 77 - 100

Questions for discussion and reflection:

1. How do I currently document/record student assessment information?
2. Why am I recording and documenting my student information in this way?
3. Am I documenting what I really value and expect in student learning?
4. How do I communicate to my students and parents what students have learned?
5. What types of communication tools do I use to give students feedback?
6. What strategies do I use to encourage students to take responsibility for their own learning?

Required activity:

- Activity 1 – Documenting Assessments and Communicating Results

Optional activities: (Choose one)

- Activity 2 – "Grading: The Issue is not How, but Why" by Alfie Kohn
- Activity 3 – Issues in Grading and Reporting
- Activity 4 – A Closer Look at *One* Assessment Topic

Checking for understanding:

Generating questions

Reflecting on what was learned:

1. How does the use of portfolios and student-led parent conferences promote greater student understanding of learning goals?

2. How does the use of rubrics and specific feedback to students promote greater student achievement?

Complete this checklist when you have finished this module.

_____ Read pp. 77-100 in the manual

_____ Discussed "Questions for discussion and reflection"

_____ Completed Activity 1 – Documenting and Communicating Results

_____ Completed "Checking for Understanding"

_____ Completed one optional activity

_____ Reflected upon what was learned

ACTIVITY 1

DOCUMENTING AND COMMUNICATING RESULTS

Directions:
Continue to build on the required activities for Chapters 2 - 4 by using that information to work through the following questions and reflections on the assessment you gave your students.

QUESTIONS TO PONDER	YOUR RESPONSE TO QUESTIONS	HOW DID YOU DO THIS? AND WHAT WOULD YOU CHANGE NEXT TIME?
Did you decide prior to giving your students this assessment **what** information would be recorded?		
Did your students understand **what** information would be recorded?		
Did you decide ahead of time **how** you would record the information you collected using this assessment?		
Did you decide ahead of time **how** the information would be communicated to the appropriate audiences?		
Was there anything unique for you about the way you documented or communicated this assessment information?		
Were you able to communicate to your students the importance of the learning targets during this assessment process?		
Did your assessment communicate to students a process for learning or did it communicate a thing done to them to see what they know?		
Did you involve the students directly in deciding when, how and why they would be assessed and how they wanted to receive their feedback?		

Activity 2

ISSUES IN GRADING AND REPORTING
"GRADING: THE ISSUE IS NOT HOW BUT WHY"

PART ONE

Write down the **three** biggest questions about grading and reporting that you hear from your colleagues.

1. _____

2. _____

3. _____

Do you feel these questions are important ones for you personally? Why?

_____

PART TWO

Read "Grading: The Issue Is Not How but Why," copyright 1994 by Alfie Kohn, reprinted from *Educational Leadership*, October 1994, with the author's permission. The article is reprinted on the following pages. Where in Kohn's three levels of concern about grading and reporting do your questions from Part I fall?

		Where do your questions fall? Why?
Level I: Level 2: Level 3:	• How to combine numbers. • Beyond traditional. • Why grade? Why assess?	

PART THREE

LIST AT LEAST ONE THING FROM THE KOHN ARTICLE IN EACH CATEGORY:
I like...
I don't like...this makes me uncomfortable...
This is interesting; I want to think more about it...

Grading: The Issue Is Not How but Why

Alfie Kohn

Only by abandoning traditional grading and performance assessment practices can we achieve our ultimate educational objectives.

Why are we concerned with evaluating how well students are doing? The question of motive, as opposed to method, can lead us to rethink basic tenets of teaching and learning, and to evaluate what students have done in a manner more consistent with our ultimate educational objectives. But not all approaches to the topic result in this sort of thoughtful reflection. In fact, approaches to assessment may be classified according to their depth of analysis and willingness to question fundamental assumptions about how and why we grade. Consider three possible levels of inquiry:

Level 1. These are the most superficial concerns, those limited to the practical issue of *how* to grade students' work. Here we find articles and books offering elaborate formulas for scoring assignments, computing points, and allocating final grades—thereby taking for granted that what students do must receive *some* grades and, by extension, that students ought to be avidly concerned about the ones they will get.

Level 2. Here educators call the above premises into question, asking whether traditional grading is really necessary or useful for assessing students' performance. Alternative assessments, often designated as "authentic," belong in this category. The idea here is to provide a richer, deeper description of students' achievement. (Portfolios of students' work are sometimes commended to us in this context, but when a portfolio is used merely as a means of arriving at a traditional grade, it might more accurately be grouped under Level 1.)

Level 3. Rather than challenging grades alone, discussions at this level challenge the whole enterprise of assessment—and specifically why we are evaluating students as opposed to *how* we are doing so. No matter how elaborate or carefully designed an assessment strategy may be, the result will not be constructive if our reason for wanting to know how students are doing is itself objectionable.

Grading Rationale I: Sorting

One reason for evaluating students is to be able to label them on the basis of their performance and thus to sort them like so many potatoes. Sorting, in turn, has been criticized at each of the three levels, but for very different reasons. At Level 1, the concern is merely that we are not correctly dumping individuals into the right piles. The major problem with our high schools and colleges, the argument goes, is that they don't keep enough students off the Excellent pile. (These critics don't put it quite this way, of course; they talk about "grade inflation.") Interestingly, most studies suggest that student performance does not improve when instructors grade more stringently and, conversely, that making it relatively easy to get a good grade does not lead students to do inferior work—even when performance is defined as the number of facts retained temporarily as measured by multiple-choice exams (Vasta and Sarmiento 1979, Abrami et al. 1980).

At Level 2, questions are raised about whether grades are reliable enough to allow students to be sorted effectively. Indeed, studies show that any particular teacher may well give different grades to a single piece of work submitted at two different times. Naturally the variation is even greater when the work is evaluated by more than one teacher (Kirschenbaum et al. 1971). What grades offer is spurious precision, a subjective rating masquerading as an objective assessment.

From the perspective of Level 3, this criticism is far too tame. The trouble is not that we are sorting students badly—a problem that logically should be addressed by trying to do it better. The trouble is that we are sorting them at all. Are we doing so in order to segregate students by ability and teach them separately? The harms of this practice have been well established (Oakes 1985). Are we turning schools into "bargain-basement personnel screening agencies for business" (Campbell 1974, p. 145)? Whatever use we make of sorting, the process itself is very different from—and often incompatible with—the goal of helping students to learn.

Grading Rationale II: Motivation

A second rationale for grading—and indeed, one of the major motives behind assessment in general—is to motivate students to work harder so they will receive a favorable evaluation. Unfortunately, this rationale is just as problematic as sorting. Indeed, given the extent to which *A*'s and *F*'s function as rewards and punishments rather than as useful feedback, grades are counterproductive regardless of whether they are intentionally used for this purpose. The trouble lies with the implicit assumption that there exists a single entity called "motivation" that

students have to a greater or lesser degree. In reality, a critical and qualitative difference exists between *intrinsic* and *extrinsic* motivation—between an interest in what one is learning for its own sake, and a mindset in which learning is viewed as a means to an end, the end being to escape a punishment or snag a reward. Not only are these two orientations distinct, but they also often pull in opposite directions.

Scores of studies in social psychology and related fields have demonstrated that extrinsic motivators frequently undermine intrinsic motivation. This may not be particularly surprising in the case of sticks, but it is no less true of carrots. People who are promised rewards for doing something tend to lose interest in whatever they had to do to obtain the reward. Studies also show that, contrary to the conventional wisdom in our society, people who have been led to think about what they will receive for engaging in a task (or for doing it well) are apt to do lower quality work than those who are not expecting to get anything at all.

These findings are consistent across a variety of subject populations, rewards, and tasks, with the most destructive effects occurring in activities that require creativity or higher-order thinking. That this effect is produced by the extrinsic motivators known as grades has been documented with students of different ages and from different cultures. Yet the findings are rarely cited by educators.

Studies have shown that the more students are induced to think about what they will get on an assignment, the more their desire to learn evaporates, and, ironically, the less well they do. Consider these findings:

- On tasks requiring varying degrees of creativity, Israeli educational psychologist Ruth Butler has repeatedly found that students perform less well and are less interested in what they are doing when being graded than when they are encouraged to focus on the task itself (Butler and Nissan 1986; Butler 1987, 1988).
- Even in the case of rote learning, students are more apt to forget what they have learned after a week or so—and are less apt to find it interesting—if they are initially advised that they will be graded on their performance (Grolnick and Ryan 1987).
- When Japanese students were told that a history test would count toward their final grade, they were less interested in the subject—and less likely to prefer tackling difficult questions—than those who were told the test was just for monitoring their progress (Kage 1991).
- Children told that they would be graded on their solution of anagrams chose easier ones to work on—and seemed to take less pleasure from solving them—than children who were not being graded (Harter 1978).

> Rather than challenging grades alone, we need to question the whole enterprise of assessment.

As an article in the *Journal of Educational Psychology* concluded, "Grades may encourage an emphasis on quantitative aspects of learning, depress creativity, foster fear of failure, and undermine interest" (Butler and Nissan 1986, p. 215). This is a particularly ironic result if the rationale for evaluating students in the first place is to encourage them to perform better.

Grading Rationale III: Feedback

Some educators insist that their purpose in evaluating students is neither to sort them nor to motivate them, but simply to provide feedback so they can learn more effectively tomorrow than they did today. From a Level 2 perspective, this is an entirely legitimate goal—and grades are an entirely inadequate means of reaching it. There is nothing wrong with helping students to internalize and work toward meeting high standards, but that is most likely to happen when they "experience success and failure not as reward and punishment, but as information" (Bruner 1961, p. 26). Grades make it very difficult to do this. Besides, reducing someone's work to a letter or number simply is not helpful; a *B+* on top of a paper tells a student nothing about what was impressive about that paper or how it could be improved.

But from Level 3 comes the following challenge: *Why do we want students to improve*? This question at first seems as simple and bland as baby food; only after a moment does it reveal a jalapeño kick: it leads us into disconcerting questions about the purpose of education itself.

Demand vs. Support

Eric Schaps (1993), who directs the Developmental Studies Center in Oakland, California, has emphasized

"a single powerful distinction: focusing on what students ought to be able to do, that is, what we will demand of them—as contrasted with focusing on what we can do to support students' development and help them learn." For lack of better labels, let us call these the "demand" and "support" models.

In the demand model, students are workers who are obligated to do a better job. Blame is leveled by saying students "chose" not to study or "earned" a certain grade—conveniently removing all responsibility from educators and deflecting attention from the curriculum and the context in which it is taught. In their evaluations, teachers report whether students did what they were supposed to do. This mind-set often lurks behind even relatively enlightened programs that emphasize performance assessment and—a common buzzword these days—*outcomes*. (It also manifests itself in the view of education as an investment, a way of preparing children to become future workers.)

> The trouble is not that we are sorting students badly, the trouble is that we are sorting them at all.

The support model, by contrast, helps children take part in an "adventure in ideas" (Nicholls and Hazzard 1993), guiding and stimulating their natural inclination to explore what is unfamiliar; to construct meaning; to develop a competence with and a passion for playing with words, numbers, and ideas. This approach meshes with what is sometimes called "learner-centered learning," in which the point is to help students act on their desire to make sense of the world. In this context, student evaluation is, in part, a way of determining how effective *we* have been as educators. In sum, improvement is not something we require of students so much as something that follows when we provide them with engaging tasks and a supportive environment.

Supportive Assessment

Here are five principles of assessment that follow from this support model:

1. Assessment of any kind should not be overdone. Getting students to become preoccupied with *how* they are doing can undermine their interest in *what* they are doing. An excessive concern with performance can erode curiosity—and, paradoxically, reduce the quality of performance. Performance-obsessed students also tend to avoid difficult tasks so they can escape a negative evaluation.

2. The best evidence we have of whether we are succeeding as educators comes from observing children's behavior rather than from test scores or grades. It comes from watching to see whether they continue arguing animatedly about an issue raised in class after the class is over, whether they come home chattering about something they discovered in school, whether they read on their own time. Where interest is sparked, skills are usually acquired. Of course, interest is difficult to quantify, but the solution is not to return to more conventional measuring methods; it is to acknowledge the limits of measurement.

3. We must transform schools into safe, caring communities. This is critical for helping students to become good learners and good people, but it is also relevant to assessment. Only in a safe place, where there is no fear of humiliation and punitive judgment, will students admit to being confused about what they have read and feel free to acknowledge their mistakes. Only by being able to ask for help will they be likely to improve.

Ironically, the climate created by an emphasis on grades, standardized testing, coercive mechanisms such as pop quizzes and compulsory recitation, and pressure on teachers to cover a prescribed curriculum makes it more difficult to know how well students understand—and thus to help them along.

4. Any responsible conversation about assessment must attend to the quality of the curriculum. The easy question is whether a student has learned something; the far more important—and unsettling—question is whether the student has been given something worth learning. (The answer to the latter question is almost certainly no if the need to evaluate students has determined curriculum content.) Research corroborates what thoughtful teachers know from experience: when students have interesting things to do, artificial inducements to boost achievement are unnecessary (Moeller and Reschke 1993).

5. Students must be invited to participate in determining the criteria by which their work will be judged, and then play a role in weighing their work against those criteria. Indeed, they should help make decisions about as many elements of their learning as possible (Kohn 1993). This achieves several things: It gives them more control over their education, makes evaluation feel less punitive, and provides an important learning experience in itself. If there is a movement away from grades, teachers should explain the rationale and solicit students' suggestions for what to do instead and how to manage the transitional period. That transition may be bumpy and slow, but the chance to engage in personal and collective reflection about these issues will be important in its own right.

And If You Must Grade ...

Finally, *while conventional grades persist, teachers and parents ought to*

do everything in their power to help students forget about them. Here are some practical suggestions for reducing the salience.

■ *Refrain from giving a letter or number grade for individual assignments,* even if you are compelled to give one at the end of the term. The data suggest that substantive comments should replace, not supplement, grades (Butler 1988). Make sure the effect of doing this is not to create suspense about what students are going to get on their report cards, which would defeat the whole purpose. Some older students may experience, especially at first, a sense of existential vertigo: a steady supply of grades has defined them. Offer to discuss privately with any such student the grade he or she would probably receive if report cards were handed out that day. With luck and skill, the requests for ratings will decrease as students come to be involved in what is being taught.

■ *Never grade students while they are still learning something and, even more important, do not reward them for their performance at that point.* Studies suggest that rewards are most destructive when given for skills still being honed (Condry and Chambers 1978). If it is unclear whether students feel ready to demonstrate what they know, there is an easy way to find out: ask them.

■ *Never grade on a curve.* The number of good grades should not be artificially limited so that one student's success makes another's less likely. Stipulating that only a few individuals can get top marks regardless of how well everyone does is egregiously unfair on its face. It also undermines collaboration and community. Of course, grades of any kind, even when they are not curved to create artificial scarcity—or deliberately publicized—tend to foster comparison and competition, an emphasis on relative standing. This is not only destructive to students' self-esteem and relationships but also counterproductive with respect to the quality of learning (Kohn 1992).

As one book on the subject puts it: "It is not a symbol of rigor to have grades fall into a 'normal' distribution; rather, it is a symbol of failure—failure to teach well, to test well, and to have any influence at all on the intellectual lives of students" (Milton et al. 1986, p. 225).

■ *Never give a separate grade for effort.* When students seem to be indifferent to what they are being asked to learn, educators sometimes respond with the very strategy that precipitated the problem in the first place—grading students' efforts to coerce them to try harder. The fatal paradox is that while coercion can sometimes elicit resentful obedience, it can never create desire. A low grade for effort is more likely to be read as "You're a failure even at trying." On the other hand, a high grade for effort combined with a low grade for achievement says "You're just too dumb to succeed." Most of all, rewarding or punishing children's efforts allows educators to ignore the possibility that the curriculum or learning environment may have something to do with students' lack of enthusiasm. ■

References

Abrami, P. C., W. J. Dickens, R. P. Perry, and L. Leventhal. (1980). "Do Teacher Standards for Assigning Grades Affect Student Evaluations of Instruction?" *Journal of Educational Psychology* 72: 107–118.

Bruner, J. S. (1961). "The Act of Discovery." *Harvard Educational Review* 31: 21–32.

Butler, R. (1987). "Task-Involving and Ego-Involving Properties of Evaluation." *Journal of Educational Psychology* 79: 474–482.

Butler, R. (1988). "Enhancing and Undermining Intrinsic Motivation." *British Journal of Educational Psychology* 58 (1988): 1–14.

Butler, R., and M. Nissan. (1986). "Effects of No Feedback, Task-Related Comments, and Grades on Intrinsic Motivation and Performance." *Journal of Educational Psychology* 78: 210–216.

Campbell, D. N. (October 1974). "On Being Number One: Competition in Education." *Phi Delta Kappan*: 143–146.

Condry, J., and J. Chambers. (1978). "Intrinsic Motivation and the Process of Learning." In *The Hidden Costs of Rewards: New Perspectives on the Psychology of Human Motivation*, edited by M. R. Lepper and D. Greene. Hillsdale, N.J.: Lawrence Erlbaum.

Grolnick, W. S., and R. M. Ryan. (1987). "Autonomy in Children's Learning: An Experimental and Individual Difference Investigation." *Journal of Personality and Social Psychology* 52: 890–898.

Harter, S. (1978). "Pleasure Derived from Challenge and the Effects of Receiving Grades on Children's Difficulty Level Choices." *Child Development* 49: 788–799.

Kage, M. (1991). "The Effects of Evaluation on Intrinsic Motivation." Paper presented at the meeting of the Japan Association of Educational Psychology, Joetsu, Japan.

Kirschenbaum, H., R. W. Napier, and S. B. Simon. (1971). *Wad-Ja-Get?: The Grading Game in American Education*. New York: Hart.

Kohn, A. (1992). *No Contest: The Case Against Competition*. Rev. ed. Boston: Houghton Mifflin.

Kohn, A. (September 1993). "Choices for Children: Why and How to Let Students Decide." *Phi Delta Kappan*: 8–20.

Milton, O., H. R. Pollio, and J. A. Eison. (1986). *Making Sense of College Grades*. San Francisco: Jossey-Bass.

Moeller, A. J., and C. Reschke. (1993). "A Second Look at Grading and Classroom Performance." *Modern Language Journal* 77: 163–169.

Nicholls, J. G., and S. P. Hazzard. (1993). *Education as Adventure: Lessons from the Second Grade*. New York: Teachers College Press.

Oakes, J. (1985). *Keeping Track: How Schools Structure Inequality*. New Haven: Yale University Press.

Schaps, E. (October 1993). Personal communication.

Vasta, R., and R. F. Sarmiento. (1979). "Liberal Grading Improves Evaluations But Not Performance." *Journal of Educational Psychology* 71: 207–211.

Alfie Kohn is the author of *Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes* (Houghton Mifflin). His address is 41 James St., Montclair, NJ 07042.

Activity 3

THINGS THAT MAY OR MAY NOT BE UNDER OUR CONTROL
TAKE TIME TO REFLECT

Directions:
Review page 99 in the manual. Decide if you agree or disagree with the advice from Rick Stiggins and reflect on why or why not.

STIGGIN'S STATEMENT	AGREE OR DISAGREE	WHY?
1. Grade on achievement of prespecified targets only, not intelligence, effort, attitude, or personality.		
2. Always rely on the most current information available about student achievement.		
3. Devise grades that reflect achievement status with respect to preset targets rather than improvement.		
4. Decide borderline cases with additional information on achievement.		
5. Keep grading procedures separate from punishment.		
6. Change all policies that lead to miscommunication about achievement.		
7. Add further detail to grade reports when needed.		
8. Expect individual accountability for learning even in cooperative environments.		
9. Give extra credit for evidence of extra learning—not just for doing extra work.		

Find a peer or faculty group to discuss these statements. How much agreement would you expect to find among your colleagues? What benefit would such discussion have for your school?

Activity 4

A Closer Look at Assessment Topics

Directions: The purpose of this optional activity is to take a closer look at student-led parent conferences, portfolios, or senior projects to determine when, why and how these are used. Choose one of the topics and respond to the questions. *(Do only one page.)*

Questions/Comments	Your Response to Student-Led Parent Conferences
1. Refer to Chapter 5, pages 90-95 and determine what a **student-led parent conference** involves.	
2. What are some of the advantages/benefits you would note for parents and students when this process is used?	
3. What are some barriers or disadvantages you would note for parents and students when this process is used?	
4. How would you work around any barriers to this process?	
5. What would be the main purpose for a teacher to utilize this assessment process?	
6. What would be your biggest question about the use of this assessment strategy? Would you ever consider using student-led parent conferences?	

Activity 4 continued

QUESTIONS/COMMENTS	YOUR RESPONSE TO PORTFOLIOS
1. Refer to Chapter 5, pages 86-89, and describe what a **portfolio** is.	
2. What are some of the advantages/benefits you would note for parents and students when this assessment process is used?	
3. What are some of the disadvantages or barriers you would note for parents and students when this assessment process is used?	
4. Would there be ways to work around the barriers to this process? How would you do that?	
5. What would be the main purpose for a teacher to utilize this assessment process?	
6. What would be your biggest question about the use of this assessment strategy? Would you every consider using portfolios?	

Activity 4 continued

QUESTIONS/COMMENTS	YOUR RESPONSE TO SENIOR PROJECTS
1. Refer to Chapter 5, pages 95 - 96, and determine what is involved with a **senior project.**	
2. What are some of the advantages/benefits you would note for parents and students when this process is used?	
3. What are some of the disadvantages or barriers you would note for parents and students when this process is used?	
4. Would there be ways to work around the barriers to this process? How would you do that?	
5. What would be the main purpose for a teacher to utilize this assessment process?	
6. What would be your biggest question about the use of this assessment strategy? Would you ever consider having your students do a senior project?	

CHECKING FOR UNDERSTANDING

Directions:
List three key questions that are important to consider for each of the following topics.

1. Deciding what information needs to be recorded
2. Recording assessment information
3. Using rubrics
4. Using portfolios
5. Using student-led parent conferences
6. Building student responsibility for learning

Final Activity – Suggested time 1 hour

Personal Reflections on Classroom Assessment

Directions:
Choose **one** question from the *Reflecting on what has been learned* section in each chapter and respond in the reflection record on the following pages.

Overview

1. How do the needs of assessment information vary for the state, district, school, teacher, parent, and student?

2. What questions or insights do I now have about assessment that I have not previously thought about?

Chapter One

3. How does thinking of assessment as an on-going process in the classroom support and enhance state testing?

4. Give examples to illustrate how assessment as a process relates to different purposes for assessment.

Chapter Two

5. How does developing a conceptual road map or knowing very clearly what I'm responsible for teaching make me a more effective teacher?

6. What are the relationships among my *Standard Course of Study*, my pacing guide, my textbook, and other resources?

Chapter Three

7. How may using multiple assessment methods result in improved student learning and achievement?

8. What is one assessment strategy that I have not used frequently and that I might consider implementing?

Chapter Four

9. How confident am I that the assessments I am using in the classroom really measure what I want to measure? Explain.

10. How do the instructional decisions that I make about individual students or groups influence their achievement at the end of the year?

Chapter Five

11. How does the use of portfolios and student-led parent conferences promote greater student understanding of learning goals?

12. How does the use of rubrics and specific feedback to students promote greater student achievement?

Reflection Record

Question from Overview
Question from Chapter One
Question from Chapter Two

Reflection Record Continued

Question from Chapter Three
Question from Chapter Four
Question from Chapter Five

Feedback for Authors

Directions: Please respond to each item, fold and staple, and mail to the address on the back.

Name _____

School _____ LEA _____

Grade/Courses _____

Please circle to indicate your opinion:	Not useful				Extremely useful
1. Was the information in the manual useful?	1	2	3	4	5
2. How useful was this study guide in helping modify your classroom assessment practices?	1	2	3	4	5
3. How useful was having the video?	1	2	3	4	5
4. How useful was having a study partner or group?	1	2	3	4	5
5. How useful were the activities in helping you understand and learn the information?	1	2	3	4	5

What suggestions do you have for improving the text and/ or the study guide?

What follow-up materials or professional development would you find helpful?

Jeane Joyner
Jan Williamson
Department of Public Instruction
301 N. Wilmington Street
Raleigh, NC 27601

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

TM029654

NOTICE

REPRODUCTION BASIS

☑ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").